

Traces of Transfer?

Pragmatic development in the use of initial adverbials in the interlanguage of advanced Dutch learners of English

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl

<http://www.lotschool.nl>

ISBN: 978-94-6093-230-4
NUR 616

Copyright © 2017: Sanne van Vuuren. All rights reserved.

Traces of Transfer?

Pragmatic development in the use of initial adverbials in
the interlanguage of advanced Dutch learners of English

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,

volgens besluit van het college van decanen

in het openbaar te verdedigen op donderdag 30 maart 2017

om 12.30 uur precies

door

Sanne van Vuuren

geboren op 10 februari 1981

te Amsterdam

Promotor: Prof. dr. A.M.C. van Kemenade

Copromotor: Dr. P.J.M. de Haan

Manuscriptcommissie:

Prof. dr. J.J.M. Schoonen

Prof. dr. M. Hannay (VU)

Prof. dr. B.L.J. Los (University of Edinburgh, Verenigd Koninkrijk)

Acknowledgements

I am very happy and mildly amazed that this thesis has emerged from the chaos of the past few years. Clearly, I have no one but myself to blame for the stresses and strains inherent in combining a PhD with having children and moving to a house in need of extensive building work. There are, however, numerous people that I have to thank for helping me rise to the occasion.

First of all, many thanks to my supervisors, Ans van Kemenade and Pieter de Haan. You have both been my teachers ever since I first stepped into the Erasmus building as a timid teenager. Over the years, I've benefited from your experience in many different ways, the most tangible result of which is this thesis. Ans, thank you for sharing your expertise and for offering me practical support and encouragement. The phrase 'onwards and upwards' will forever remind me of you. Pieter, thank you for your tireless work in compiling the Dutch component of the LONGDALE corpus, for introducing me to your network of learner corpus researchers and for involving me in the organisation of the LCR2015 conference, which has been an enormously rewarding experience.

Thank you to my room-mates Gea, Erwin, Rosanne and, last but not least, Meta. As work has been a big part of my life in these last few years, you have been a big part of it as well. Thank you for being great company. Erwin, I owe you a second thank you for your never-ending patience in helping me understand the intricacies of Corpus Studio and Cesax. The software you developed (and adapted, if the need arose) has been an essential ingredient in this project.

Thank you very much also to all of my other colleagues. It's wonderful to work together with people who share a love of language in one form or another. A special thank you to our former colleague, Bettelou Los, whose ideas inspired me to look into the issues this thesis explores. Bettelou, thank you for your endless enthusiasm. You're a great motivator.

A warm thank you to my co-authors, Janine, Lyuben and Rina. You have each brought your own approaches and specialist knowledge to the articles we collaborated on. I have learned a lot from working with you. I look forward to many more exciting projects that allow us to join forces and share endless cups of coffee. Rina, thank you also for your extraordinary generosity in sharing your materials, for spending your time helping me get to grips with the ESP course you had invested so much in, and for your tremendous help in proofreading my thesis.

Thank you to the editors and anonymous reviewers of *Linguistics in the Netherlands* and the *International Journal of Learner Corpus Research*, the participants in the LiTS talks and the 2013 and 2015 Learner Corpus Research conferences, the international LONGDALE team, my manuscript committee, as

well as the many other people who have either formally or informally provided me with valuable feedback on presentations and earlier drafts of the papers combined into this volume.

Thank you to my mother, Mieke, for leaving your wonderful mountain retreat in France and coming to stay with us for weeks on end if we needed your help to look after the children when Oleg was abroad. You always leave us much more organised than you find us, using your time to not only look after the children, but also to paint wardrobes, install shelves, sew curtains and tame our garden. Thank you for everything.

Thank you to my father, Leen, and Anna, for welcoming the children whenever I need to go to conferences. At times, reading a book in a child-free hotel room feels like one of the greatest luxuries imaginable, but I don't think I would be able to enjoy it at all if I did not know that the children were having the time of their life with you. Of course, you haven't just looked after the children, but after us as well. If it wasn't for your help, our house would be much less inhabitable, lacking floors, doors or walls in practically every room in the house.

Thank you too to the rest of my family, not least to Gertje and Hans, for quite literally buying me time when I needed it most. Thank you, Maja, for being so much more than a kind, warm-hearted and reliable babysitter. Without your unflappable presence, I'm not sure I would have managed to complete this project at all.

Thank you to my incredibly thoughtful, intelligent and generous friends, Wendelien, Mihriban and Mylène in particular. Your optimism and laughter are absolutely irresistible. The fact that I have managed to retain my sanity in these past few years is in no small measure thanks to you.

Thank you to my one and only Oleg, unconventional and amazing in equal measure. Thank you for your music, for looking after me, for always sneaking cakes into my bag (and thank you to Meta and any other colleagues who happened to be around for helping me eat them. I would not have managed without you!). May we always combine the best of both worlds. Дякую за твою турботу і твою любов. Я кохаю тебе.

Dearest Luka and Ilya, you are the best children I could ever have wished for: sweet, strong and very patient. I owe you several weekends, two holidays and at least three photo albums. However much I love my work, I love you more.

Contents

1. Introduction	13
Corpus selection and annotation	21
Thesis outline.....	30
2. Information-structural transfer in advanced Dutch EFL writing: A cross- linguistic longitudinal study	33
Introduction.....	34
Comparative pragmatics of first position in Dutch and English	35
Method	40
Corpus design	40
Procedure.....	41
Results and discussion	41
Text types	41
Place.....	41
Addition.....	46
Conclusion	50
3. Common framework, local context, local anchors: How information- structural transfer can help to distinguish within <i>CEFR C2</i>	53
Introduction.....	54
<i>The Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)</i>	54
The <i>CEFR</i> in the Dutch educational system.....	56

L1 transfer at higher levels of acquisition	60
Materials and methods	62
Corpus design	62
Procedure.....	63
Normalisation.....	65
Statistical analysis	65
Results	66
Native speakers and Dutch students of English compared ..	66
OOPT as a predictor of adverbial use	70
Discussion	71
Conclusion	74

4. Dutch learner English in close-up: A Bayesian corpus analysis of pre-subject adverbials in advanced Dutch EFL writing	75
Introduction.....	76
Materials and methods	80
Corpus design and annotation	80
Data analysis.....	82
Motivation for using Bayesian analysis	82
Bayesian parameter estimation	83
Drop-outs	86
Main analysis	87
Results	88
Text type, discipline and writing expertise in NS writing...	88
Adverbials overall and circumstance, linking and stance adverbials.....	88
Local anchors	92

Argumentative and literature essays in the Dutch component of LONGDALE	95
Adverbials overall and circumstance, linking and stance adverbials	95
Local anchors	97
Novice NNS and expert NS literature papers:	
LONGDALE-lit vs VUNSPRAC-lit.....	98
Adverbials overall and circumstance, linking and stance adverbials.....	98
Local anchors	104
Argumentative writing by novice NS and NNS writers:	
LONGDALE-arg vs LOCNESS	107
Adverbials overall and circumstance, linking and stance adverbials.....	107
Local anchors	109
Timed vs untimed NNS writing	112
Adverbials overall and circumstance, linking and stance adverbials.....	112
Local anchors	115
Discussion.....	116
Conclusion.....	120

5. Instant cohesion: Exploring the role of transfer and teaching in the use of cohesive adverbials in L2 English writing	123
Introduction.....	124
Methodology.....	126
Corpus construction and annotation	126

Statistical analysis.....	127
Corpus findings.....	128
The role of teaching	134
English coursebooks in Dutch schools	134
Occurrence of linking devices in coursebooks.....	135
Presentation with or without context	136
Instructions for usage	136
Exercises.....	137
Alternative cohesive strategies	140
Final secondary-school reading exams	141
Teaching transfer	145
Discussion and conclusion	147

6. Same difference? L1 influence in the use of initial adverbials in English

novice writing	153
Introduction.....	154
Methodology.....	157
Corpora	157
Statistical analysis.....	160
Qualitative results: frequencies of occurrence	160
Adverbials overall	160
Circumstance adverbials	164
Stance adverbials	167
Linking adverbials.....	169
Qualitative results: adverbials in context	172
Circumstance adverbials	172
Time	174

Place	177
Adverbial length	181
Stance adverbials.....	184
Viewpoint or perspective	186
Actuality and reality.....	189
Linking adverbials	191
Addition	193
Contrast/concession.....	195
Stylistic coherence.....	197
Discussion/conclusion	198
7. Conclusion	203
Summary	203
Implications	210
Future directions	213
References	219
Appendix I Overview of circumstance, stance and linking adverbials in English coursebooks.....	231
Appendix II Adverbial categories represented in English argumentative essays by L1 Dutch, L1 French and NS novice writers.....	241
Samenvatting (Summary in Dutch)	243
Curriculum Vitae	255

12 - *Traces of Transfer*

1. Introduction

Slowly but surely, mass bilingualism is gaining ground in the Netherlands. Although English is not an official language, a recent large-scale survey of uses of and attitudes to English in the Netherlands reveals that 90% of respondents 'rated themselves as having at least reasonable to fluent reading, writing, speaking and listening skills' (Edwards, 2016, p. 97). This confirms the conclusions of an earlier EU survey, which also reported that, at 90% of the population, the Netherlands has the greatest proportion of non-native speakers (NNSs) of English in Europe (Eurobarometer, 2012). Driven by globalisation and facilitated by education, competence in English has come to be regarded as 'a basic skill universally acquired' (Edwards, 2016, p. 197). Nowhere can this be seen more clearly than at universities across the country, where the wholesale switch to English as the main language of instruction for most master and an increasing number of bachelor programmes has made English indispensable for staff and students alike.

But there are also students for whom English is not just a means but an object of study. For a combination of affective and practical reasons, these future language professionals – teachers, translators, ESP trainers, coursebook authors, editors – are likely to set their sights beyond communicative competence (Granger, 2004; De Haan & Van der Haagen, 2012; Flowerdew, 2015). It is the challenge faced by these very advanced language learners that this thesis will be concerned with. The available literature on learners at comparably advanced stages of acquisition suggests that they have 'typically mastered the lexicogrammatical rules of English morphosyntax' (Springer, 2012, p. 14), that their language is characterised by 'a relatively high degree of formal accuracy' (Carroll *et al.*, 2000, p. 461) and is 'mainly free from grave grammatical errors' (Callies,

2009). In short, Carroll and Lambert (2003) assert, 'the learning problem at advanced stages of learning is not one of linguistic form' (270). Learners may, however, not be fully aware of pragmatic considerations in the choice between formally correct alternatives. In the words of Bardovi-Harlig (1999), 'high levels of grammatical competence do not guarantee concomitant high levels of pragmatic competence' (p. 686).

Learners' pragmatic competence has been the subject of extensive research in the field of interlanguage pragmatics. While the field has traditionally been mainly concerned with intercultural production and understanding of speech acts (Bardovi-Harlig, 1999; Callies, 2007, 2009; Barron, 2011), Callies (2007) proposes an extension of its scope by focusing on learners' pragmalinguistic (as opposed to sociopragmatic) choices in written text. He defines pragmalinguistic knowledge as 'a component of L2 pragmatic knowledge which relates to learners' knowledge of the structural linguistic resources available in a given language for realizing particular communicative effects, and knowledge of the appropriate contextual use of these resources' (p. 12). This reorientation of the discipline of interlanguage pragmatics as well as related research in the fields of psycholinguistics and translation studies has revealed that the interlanguage of very advanced learners of typologically related languages is in many cases characterised by transfer in the pragmalinguistic subdomain of information structuring (e.g. Hannay & Keizer, 1993; Carroll *et al.*, 2000; Carroll & Lambert, 2003, 2006; Bohnacker & Rosén, 2007, 2008; Callies, 2009; Dimroth & Starren, 2003). The interpretation of the type of evidence used to make the case for information-structural transfer is not always straightforward, however. While frequential differences in the use of certain syntactic constructions may point to underlying information-structural differences, Bohnacker and Rosén (2007)

concede that 'it may be difficult to distinguish between transfer of information structure and transfer of frequencies of syntactic constructions' (p. 54).

Apart from the nature of transfer at advanced stages of acquisition, questions also remain about its directionality. Jarvis (2000) reviews various studies, some of which apparently demonstrate that L1 transfer decreases with proficiency, others that it increases, remains constant or fluctuates (p. 247). He argues that these inconsistencies may be attributed to the lack of a unified framework for studying transfer and goes on to consider the types of evidence that any transfer study should minimally consider in order to ensure that 'results are mutually comparable and findings can truly be generalised' (Jarvis, 2000, p. 248):

- 1) intra-L1-group homogeneity in learners' interlanguage performance
- 2) inter-L1-group heterogeneity in learners' interlanguage performance
- 3) intra-L1-group congruity between learners' L1 and interlanguage performance

The first type of evidence serves to demonstrate that learners with the same L1-background behave alike, ruling out genre-based or individual variation. The second type of evidence, inter-L1-group heterogeneity, reflects the differences between groups of learners with different L1 backgrounds, which 'strengthens the argument for L1 influence because it essentially rules out developmental and universal factors as the cause of the observed interlanguage behaviour' (Jarvis, 2000, p. 254-5). Intra-L1-group congruity between learners' L1 and interlanguage performance, finally, confirms transfer as the likely cause of an observed linguistic pattern as it can be observed both in learners' L1 and in their interlanguage productions.

In considering the remaining linguistic challenges faced by very advanced Dutch learners of English, this study will focus attention on a feature at the interface of syntax and information structure: the use of clause-initial, pre-subject adverbials. Previous research suggests that advanced Dutch learners of English tend to begin their clauses with adverbial phrases more frequently than native speakers (NSs) do and that they may not recognise contextual restrictions on placement of adverbials in initial position, often leading to productions which have aptly been characterised as ‘top-heavy’ or suffering from ‘frontal overload’ (Hannay & Keizer, 1993, p. 20), as in the following example:

- 1) In a number of districts an investigation into the consequences of the new coffee shop policy is being carried out.

(Hannay & Mackenzie, 2009, p. 127)

Frontal overload – and the corresponding informationally light sentence ending – does not depend on adverbial placement only, but rather results from the combination of an initial adverbial, a focal subject and a passive, which effectively cluster all the informational content in front of the verb, leaving the end of the sentence to peter out. The placement of the adverbial in pre-subject position, however, does affect its interpretation. In English, ‘[t]here is a real danger that the reader will interpret “in a number of districts” as the most important information in the sentence’ (Hannay & Mackenzie, 2009, p. 127), while its Dutch equivalent can function as a neutral point of departure.

This pragmatic distinction has been argued to derive from the typological differences between both languages. As a subject-verb-object (SVO) language, English has a relatively inflexible sentence structure, unlike a verb-second (V2)

language like Dutch, in which the verb is fixed in second position in declarative main clauses, but other constituents can occur either pre- or post-verbally:

Table 1. V2 word order

	X	V-finite	Subject	Object	Adverbial
a.	Joris	zag		gouden eieren	in het drakenhol
b.	In het drakenhol	zag	Joris	gouden eieren	
c.	Gouden eieren	zag	Joris		in het drakenhol
George saw golden eggs in the dragon's den.					

There is both diachronic and synchronic evidence that differences in the pragmatic function of a preverbal clause-initial constituent in a V2 language like Dutch and a pre-subject clause-initial constituent in an SVO language like English may affect the markedness of adverbials such as in (1) above. Los (2009; 2012) and Los and Dreschler (2012) have shown that the transition of English from V2 to SVO syntax has affected the functionality of the first position:

‘In verb-second, a single constituent from anywhere in the clause is placed before the finite verb. This constituent is very versatile with respect to its discourse functions: it may provide marked or unmarked links with the preceding discourse. After verb-second is lost, [the] discourse functions of the first constituent appear to be allotted to different positions: presubject position, which is pragmatically marked, and subject position, which is pragmatically neutral.’

(Los, 2009, p. 118)

18 - *Traces of Transfer*

In the relatively inflexible SVO structure of present-day English, pre-subject constituents are optional and, if they are included at all, are more likely to be perceived as marked, i.e. to be interpreted as particularly important, possibly even contrastive, as in example (1) above. The preverbal slot in a V2 system, on the other hand, has to be filled and may host marked or unmarked subjects, objects or adverbials. This implies that the placement of the grammatical subject is allowed to vary with its information status. If it represents new information, it can easily occur post-verbally, leaving the preverbal slot free for discourse linking. Los and Dreschler (2012) conclude that ‘Dutch and German may use the first position made available in their V2 syntax for adverbials to encode local anchors’, which they define as ‘explicit links to the preceding discourse’ in the form of adverbials which ‘typically contain possessives or demonstratives’ (p. 860, p. 862). Examples include pronominal adverbs such as Dutch *daarbij* (with that) and *hiernaast* (next to this), but also (mostly locative) prepositional phrases such as the following:

2) *Een bekend heldenverhaal is het sprookje van*
A well-known heroic tale is the fairytale of

Joris en de draak. In dit sprookje vecht Joris
George and the dragon. In this fairytale fights George

met een afschuwelijke draak om een dorpje van
with a terrible dragon to a village from

de ondergang te redden.
the destruction to save.

While the use of the initial adverbial for local anchoring (in bold) is clearly not ruled out by SVO syntax, Dutch-English translation manuals, such as Lemmens and Parr (1995), warn against literal translations of initial prepositional phrases, advising translators to ‘turn the constituent following the Dutch preposition into the subject of the corresponding English sentence’ (p. 92, my translation).

If the loss of verb-second caused the grammatical subject to come to be regarded as the default option for neutral discourse linking, this may be seen as a motivation for the versatility of subjects in present-day English, with its ready acceptance of permissive subjects (Los, 2009, p. 118). Dreschler and Hebing (2010) report the ratio between subject-initial sentences and sentences starting with other first constituents as 77% against 23% respectively for English, compared with 54% against 46% for Dutch (p. 63). This SVO-V2 typological difference is reflected in the interlanguage of Danish and Norwegian learners of English, who have been shown to use more initial adverbials than native speakers (Shaw, 2004; Hasselgård, 2009). An interesting case is presented by a similar kind of transfer effect, but in the opposite direction, in texts produced by advanced Swedish learners of German (Rosén, 2006; Bohnacker & Rosén, 2007, 2008). Swedish learners have been found to start more of their sentences with subjects and fewer with other preverbal constituents than native speakers of German. However, as Swedish and German are both V2 languages, this implies that V2 syntax alone does not automatically lead speakers to use more initial adverbials and other non-subject constituents. Rather, Los (2009) argues, it is V2 syntax combining with ‘a system of deictic referencing [that is] sufficiently articulated’ that ensures the discourse linking function of the preverbal slot (p. 106).

Like Danish and Norwegian learners, Dutch learners have also been shown to use more initial adverbials than NSs of comparable age and academic background (Verheijen *et al.*, 2013). In an exploratory analysis of 137 essays,

Verheijen *et al.* find ‘an overall pattern of decrease’ in the frequency of initial adverbials (excluding linking adverbials) and pronominal adverbs in the written English of Dutch students of English Language and Culture between year one and two of their studies. They conclude, therefore, that ‘this feature ... is a good marker of progress in advanced English’ (p. 103). What remains unclear is whether this conclusion might have been supported by statistical analysis, which they note was not attempted due to the problem of attrition in longitudinal data, and how much of students’ progress (and the fluctuation by which it is accompanied) may be down to genre differences between descriptive, informative and argumentative writing.

The current study develops this research line in detail, considering the use of initial adverbials in argumentative essays and literature essays written by four cohorts of Dutch students of English Language and Culture between year 1 and year 3 of their studies, distinguishing not only between different text types but also between timed and untimed productions, in relation to both novice and expert NS writing. Like Verheijen *et al.* (2013), it departs from the hypothesis that information-structural transfer from Dutch will lead L1 Dutch learners of English to use more initial adverbials than NSs and that this hypothesised transfer effect will decrease with increasing proficiency. Apart from offering a descriptive contrastive and developmental account of the way in which Dutch learners at various stages of interlanguage development differ from NSs in the frequencies and pragmatic use of initial adverbials, this thesis also aims to explicitly look into the possible underlying causes of Dutch learners’ linguistic behaviour, i.e. evaluate the evidence for a transfer-based explanation and look into the role of teaching. This is reflected in the research questions this thesis aims to answer:

- 1) In what way do advanced learners differ from NSs in the frequency and realisation of clause-initial adverbials?
- 2) How do a) proficiency level and b) longitudinal development affect Dutch learners' use of initial adverbials?
- 3) Is Dutch learners' use of initial adverbials likely to be transfer-induced?
 - a. Is there intra-L1-group homogeneity in their interlanguage performance?
 - b. Is there inter-L1-group heterogeneity between L1 Dutch and L1 French learners?
 - c. Is there intra-L1-group congruity between their L1 and interlanguage performance?
- 4) Is Dutch learners' use of initial adverbials likely to be teaching-induced?

The framework within which the present study will try to answer these questions is a combination of Contrastive Analysis of L1 Dutch and L1 English and Contrastive Interlanguage Analysis (Granger, 1996) of L2 English by L1 Dutch learners in relation to L1 English, as well as of L2 English by L1 Dutch and Francophone learners. These two perspectives are integrated in a comparison of L1 Dutch and L2 English by L1 Dutch learners. Granger (2004) has referred to this type of research design as an 'integrated contrastive perspective, ... a very reliable empirical platform from which to conduct interlanguage research' (p. 18).

1.1 Corpus selection and annotation

While this study builds on Verheijen *et al.* (2013), it adopts a new methodology necessitated by its wider scope. It presents a corpus analysis of 1,026,056 words of text in total, which can be broken down into five different collections of L1 and L2 writing, as follows:

Table 2. sub-corpora

L1 English novice writing	59,229 words
L1 English expert writing	150,011 words
L2 English novice writing by advanced L1 Dutch learners	619,782 words
L2 English novice writing by advanced L1 French learners	138,343 words
L1 Dutch novice writing	58,691 words

The Dutch EFL essays are part of LONGDALE, the longitudinal database of learner English (Meunier, 2015), and consist of the productions of four cohorts of Dutch students of English Language and Culture at Radboud University, Nijmegen, who started their studies between 2008 and 2011. One additional set of essays, produced by cohort 2013, was aligned with the *Common European Framework of Reference (CEFR)* by means of an Oxford Online Placement Test (OOPT). The Francophone EFL essays constitute another sub-corpus of the same longitudinal database of learner English and were collected at the Université Catholique de Louvain, Belgium. Philip Springer's VUNSPRAC, the VU Native English Research Article Corpus (Springer, 2012), was used to represent L1 English expert writing, while the source of the texts produced by L1 English novice writers is LOCNESS, the Louvain Corpus of Native English Essays (Granger, 1996). The L1 Dutch novice texts, finally, were collected at Radboud University Nijmegen in September 2015.

The choice to include both novice and expert NS writing was based on the recognition that the former can perhaps more realistically be used as a reference point, whereas the latter would provide a more accurate model. Lorenz (1999) has argued that comparing L2 novice writing with L1 novice writing is both fairer and more descriptively adequate than comparing L2 novice writing with L1 expert writing. Others have noted that expert writing may represent a more relevant model for advanced L2 learners (Gilquin *et al.*, 2007; Springer, 2012). Indeed, if academic writing is 'not part of the native speaker's inheritance' but

‘acquired [...] through lengthy formal education’ (Ferguson *et al.*, 2011, p. 42), it is clear that NS expert writing may serve as a model for L1 and L2 novice writers alike. As it has been argued that frequent use of initial linking adverbials is a feature of novice writing in general, not limited to those for whom English is an L2 (Leńko-Szymańska, 2008), using an NS novice corpus along with an NS expert corpus serves to establish to what extent initial adverbial use may be determined by developmental or cognitive factors (cf. Springer, 2012). As Guilquin *et al.* point out, ‘[the] issue of the degree of overlap between novice native writers and non-native writers has far-reaching methodological and pedagogical implications and is clearly in need of empirical studies’ (2007, p. 3).

Similarly, the inclusion of a Francophone English corpus is motivated by the need to establish the heterogeneity of L2 writers with different L1s before concluding that a particular interlanguage feature is transfer-induced (cf. Jarvis, 2000). French was chosen as it is not just a different language, but also a typologically different language, Romance SVO as opposed to Germanic V2. Psycholinguistic research by Carroll and Lambert (2003, 2006) has pointed to the similarity of the Romance languages and English, not only in the fact that the position of the syntactic subject is fixed in front of the verb, but also in the preferences for information selection this entails, as ‘categories of information which have a high potential in assuming the role of topic are those which map into the syntactic subject’ (p. 269). In V2 languages such as Dutch and German, on the other hand, the syntactic subject can occur post-verbally, so that ‘information and constituents (adverbials or prepositional phrases) with topic status [...] can then systematically fill the preverbal ‘slot’ or “Vorfeld”’ (p. 269). A corpus comparison of L2 English written by advanced Dutch and Francophone learners will therefore serve to confirm or disconfirm the role of typological transfer of information-structural preferences.

The inclusion of these different sub-corpora as well as the relatively large word count of the L2 corpora in particular is of course good news in terms of representativeness and generalisability (cf. Granger, 2004, p. 3-4). It does, however, present a problem in terms of methodology, as manually sifting out clause-initial adverbials in a corpus this size is simply not feasible. Granger (2004) remarks that ‘the benefit that researchers can derive from automating some of their work is so great that it would seem a pity to do without the invaluable help it can provide’ (p. 6). Mostly, these automated approaches involve concordancing or part-of-speech tagging, but these are of limited use for a study that aims to consider a feature at the interface of syntax and information structure. The combination of corpus size and the nature of the feature under investigation therefore directed this project into relatively unexplored territory: that of automated parsing of learner corpora (cf. De Mönnink, 2000; Meunier, 2015). While a parser trained on native-speaker corpora cannot be expected to perform 100% accurately on learner data, the combined use of the Stanford Parser (Klein & Manning, 2003) and the corpus-analysis tool Corpus Studio (Komen, 2012) resulted in correct identification of between 89.55% and 96.55% of all initial adverbials occurring in declarative main clauses in a representative sample of 15,840 words. Adverbials that were not identified correctly were in many cases preceded by a comma splice or a dash or included in a main clause between brackets. The variation in the rate of correct identification of initial adverbials can therefore partly be explained by the fact that NNS writing and, to a certain extent, NS novice writing, tends to feature more unorthodox punctuation. As any false positives were removed manually, the frequencies reported on in this thesis are a conservative estimate of the true ratio of occurrence.

Although the limited accuracy of automated parsing may represent a drawback, there are considerable benefits associated with this procedure as well.

While other large-scale corpus studies have adopted a lexical approach, identifying relevant adverbials by means of standard grammars or previous research and proceeding to quantify these in the corpus under investigation, the analyses reported on in this thesis are not limited to a finite set of previously identified adverbials. This means, first of all, that it is possible to retrieve various nonstandard alternatives, including misspellings, which are probably more prevalent in novice writing in general and NNS novice writing in particular so that overlooking them would amount to underestimating the true ratio of occurrence of initial adverbials in text produced by these specific groups of writers. Secondly, and more importantly for the purpose of this study, automated parsing makes it possible to look beyond the well-researched class of linking adverbials (e.g. Milton & Tsang, 1993; Field, 1993; Granger & Tyson, 1996; Altenberg & Tapper, 1998) and include circumstance and stance adverbials as well, neither of which could possibly be represented by a finite list (cf. Biber *et al.*, 1999).

To take account of the various pragmatic functions and positional preferences of different categories of adverbials, the database of initial adverbial phrases that resulted from the automated parsing procedure and subsequent Corpus Studio query was further manually annotated with semantic category labels using the Cesax corpus-annotation tool (Komen, 2012). These labels were derived from the classification used in the comprehensive corpus-based Longman grammar (Biber *et al.*, 1999):

Table 3. Semantic categories per adverbial class according to Biber *et al.* (1999)

Circumstance	Stance	Linking
<input type="checkbox"/> time <ul style="list-style-type: none"> • position in time • duration • frequency • temporal relationship 	<input type="checkbox"/> epistemic <ul style="list-style-type: none"> • doubt and certainty • actuality and reality • source of knowledge • limitation • viewpoint or perspective • imprecision 	<input type="checkbox"/> enumeration and addition
<input type="checkbox"/> place <ul style="list-style-type: none"> • distance • direction • position 	<input type="checkbox"/> attitude	<input type="checkbox"/> summation
<input type="checkbox"/> process <ul style="list-style-type: none"> • manner • comparison • means • instrument 	<input type="checkbox"/> style	<input type="checkbox"/> apposition
<input type="checkbox"/> contingency <ul style="list-style-type: none"> • cause/reason • purpose • concession • condition • result 		<input type="checkbox"/> result/inference
<input type="checkbox"/> extent/degree <ul style="list-style-type: none"> • amplifiers • diminishers 		<input type="checkbox"/> contrast/concession
<input type="checkbox"/> addition/restriction <ul style="list-style-type: none"> • additive • restrictive 		<input type="checkbox"/> transition
<input type="checkbox"/> recipient		
<input type="checkbox"/> other		

One complicating factor in the annotation procedure is the degree of overlap between different categories of adverbials. For example, many circumstance

adverbials, time adverbials in particular, are similar to linking adverbials in that they also have a cohesive function. Nevertheless, they were categorised as circumstance adverbials, as ‘the connective function is made semantically, through the circumstantial information which indicates time relationships’ (Biber *et al.*, 1999, p. 783). Similarly, within the class of circumstance adverbials there is considerable overlap between the categories of place and time. Adverbials such as ‘in primary school’ or ‘in chapter 3’ were consistently categorised as place adverbials, even though they contain an implicit time reference as well. Other adverbials that ‘fit primarily into one category, but have secondary roles that fit another semantic category’ (782) were also categorised in accordance with Biber *et al.*’s analysis of their primary semantic function.

In Biber *et al.*’s original classification the category ‘other’ (cf. Table 3 above) includes adverbials that ‘serve in some way to show in what **respect** the action or state described in the clause is relevant or true’ (Biber *et al.*, 1999, p. 781). Although most of the adverbials Biber *et al.* include in this category typically occur in final position, there are also adverbials that fit this description that may occur initially and are therefore relevant in the context of this thesis, such as ‘with respect to’, ‘on the subject of’ and ‘as for’. In line with Mittwoch *et al.* (2002), the annotation scheme adopted in this thesis includes these topic markers in a separate category, labelled ‘domain’, as they serve to ‘restrict the domain to which the rest of the clause applies’ (p. 765).

Each adverbial was also assigned one of five referential state category labels (Figure 1), i.e. ‘new’, for those that introduce a new referent into the text, ‘inert’ for adverbials that do not refer to either a textual or non-textual referent, ‘assumed’ for adverbials with a non-textual referent that is assumed to be part of the reader’s world knowledge, ‘inferred’ for those that can be inferred from a

textual antecedent, and ‘identity’ for adverbials that refer to the same referent as a textual antecedent (Komen, 2013):

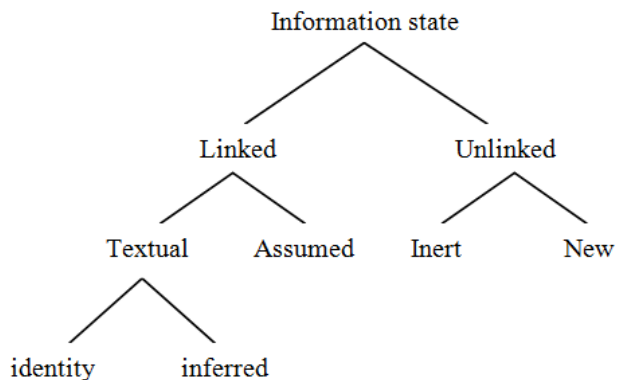


Figure 1. Referential state categories (Komen, 2013)

For those adverbials with a textual link, the referential state label also indicates the distance between the adverbial and its antecedent.

The richly annotated database that resulted from this procedure may hopefully provide a useful resource for future studies, for example by allowing a comparison of the ratio between linked and unlinked place adverbials or by specifically focusing on the use of domain adverbials as stage-setting devices. For the purpose of this thesis, however, the distinctions that were considered relevant were those between the classes of circumstance, stance and linking adverbials, as they are associated with different preferences for placement in initial, medial or final position (Biber *et al.*, 1999). Also, in line with Los and Dreschler’s (2012) association between V2 syntax and ‘the function of first constituent adverbials as local anchors’, a distinction was made between overall frequencies of adverbials belonging to each of the three main adverbial classes and frequencies of those that

were used for local anchoring, i.e. that had an identity link to the directly preceding context. The resulting classification is exemplified¹ in Table 2 below:

Table 4. Adverbial annotation scheme

adverbial class (and sub-category)	local anchor	example
circumstance (place)	no	This means I have to do all sorts of things I am not used to [...]. At home my parents would take care of these things, but I guess I will manage. (student RAD0922, year 1)
circumstance (time)	yes	<u>The Harlem Renaissance</u> , also known as the New Negro Movement, took place during the twenties and thirties of the previous century. Before this renaissance a lot of colored people had moved to the Northern regions of the United States, away for the Southern regions, because of more job opportunities becoming available in the North and because of the restraints on colored people, which were still present in the South. (student RAD0802, year 1)
stance (attitude)	no	Orfeo leaves his kingdom in the hands of his steward and wanders in the forest for years until he meets her again. Unfortunately they cannot talk to each other and Orfeo decides to follow her wherever she may go. (student RAD0813, year 2)
linking (contrast/concession)	no	Once she has become “a little colored girl” instead of just a human being, it is not hard to feel sorry for her. However , she clearly is not looking for sympathy, she is a very proud and strong woman. (student RAD0803, year 1)
linking (addition)	yes	Besides, <u>Connie makes her own decisions, namely, the fact that she decides to have love affairs, wants to have a child, and goes to Venice.</u> In addition to this , the novel is radical because of the gender relationships across class boundaries. (student RAD0927, year 2)

¹ Throughout this thesis, all examples taken from the corpora are reproduced verbatim, including any infelicities.

1.2 Thesis outline

This thesis is structured as follows: chapter 2 presents a case study on the role of L1 transfer of language-specific features of information structure in very advanced L2 learners. It contributes to answering research questions 1 and 2 by means of a contrastive and longitudinal analysis of Dutch EFL learners' use of initial place and addition adverbials in general and those that are used for local anchoring in particular.

Chapter 3 addresses the problem of defining linguistic development at advanced stages of L2 acquisition. It approaches research question 2 from a different angle, focusing not on longitudinal development but on the correlation between proficiency level and use of local anchoring in a cross-sectional *CEFR*-aligned corpus of Dutch EFL writing. In doing so, it aims to determine whether local anchoring may be considered as a possible means of differentiation within *CEFR* level C2.

The comprehensive longitudinal Bayesian analysis of initial adverbial use reported on in chapter 4 extends the contrastive-longitudinal analysis of chapter 2 and serves to make up for the lack of longitudinal statistical testing in previous studies. It assesses the development of L2 English novice writing by advanced Dutch learners and compares it with L1 English novice as well as expert writing. It also considers the role of text type, academic discipline, as well as differences between timed and untimed writing.

Chapter 5 proceeds to consider the underlying causes of Dutch learners' use of initial adverbials, investigating the extent to which it may be either teaching- or transfer-induced. It tests the source of Dutch learners' hypothesised information-structural transfer by means of a corpus analysis of L1 Dutch novice writing and L2 English novice writing by advanced Dutch learners of English, in relation to L1 English novice writing. This analysis is complemented by a

discussion of the way text structure and cohesion are treated in English and Dutch coursebooks in pre-university education.

Chapter 6 further explores the role of transfer by considering the use of clause-initial adverbials in novice writing produced by Dutch and Francophone learners of English in relation to NS novice writing. This serves to determine whether frequent use of initial adverbials may be a V2 transfer effect or represents an interlanguage feature shared by learners whose L1 is a romance SVO language or even a feature of novice writing in general, shared by NS and NNS writers alike.

Chapter 7, finally, draws the different strands of this thesis together by summarising what it is about initial adverbial use that may be said to be characteristic of Dutch English and reviewing the evidence for its two likely causes: teaching and transfer. It also considers future directions for both research and teaching practice.

2. Information-structural transfer in advanced Dutch EFL writing: a cross-linguistic longitudinal study²

Abstract

This article presents a case study on the role of L1 transfer of language-specific features of information structure in very advanced L2 learners. Cross-linguistic differences in the information status of clause-initial position in a V2 language like Dutch compared with an SVO language like English are hypothesised to result in overuse of clause-initial adverbials in the writing of advanced Dutch learners of English. This hypothesis was tested by evaluating advanced Dutch EFL learners' use of clause-initial place and addition adverbials in a syntactically annotated longitudinal corpus of student writing, compared with a native reference corpus. Results indicate that Dutch EFL learners overuse clause-initial place and addition adverbials that refer back to an antecedent in the directly preceding discourse. Although there is a clear development in the direction of native writing, transfer of information-structural features of Dutch can still be observed even after three years of extended academic exposure.

Keywords: L2 acquisition, information structure, L1 transfer, advanced learners, EFL, clause-initial position

² This is a revised version of: Van Vuuren, S. (2013). Information structural transfer in advanced Dutch EFL writing: A cross-linguistic longitudinal study. *Linguistics in the Netherlands*, 30(1), 173-187.

1. Introduction

Dutch students of English Language and Culture are expected to reach a near-native level of proficiency, and indeed their writing tends to be relatively error-free (Springer, 2012; De Haan & Van der Haagen, 2012). Then why is it that even at these advanced stages of acquisition their writing is often considered to be recognisably Dutch? As most language acquisition research has focused on earlier stages of L2 acquisition, relatively little is known about advanced learners, but it is clear that, as Carroll and Lambert (2003, p. 270) have noted, ‘the learning problem at advanced stages of learning is not one of linguistic form’. Rather, it appears that advanced learners differ from native speakers in (1) the frequency with which they use lexico-grammatical devices and syntactic structures available in the language (cf. Springer, 2012; De Haan & Van der Haagen, 2012) and (2) in the application of language-specific principles of information structure (Bohnacker & Rosén, 2008; Callies, 2009; Verheijen *et al.*, 2013). Callies, for example, notes that

[e]ven at advanced stages of L2 acquisition, in which ILs can be considered near-native in many respects, some core principles of information structure typical of those found in the native languages of the learners are retained, and divergences can be attributed to fundamental principles of organization underlying information structure.

(Callies, 2009, p. 104)

Similarly, following Verheijen *et al.* (2013), we hypothesise that the appearance of non-nativeness of texts written by advanced Dutch learners of English is largely due to the frequency with which these learners use certain types of clause-initial adverbials in their L2. This in turn might be attributed to transfer caused by an

interaction of syntactic and information-structural differences in the use of clause-initial constituents between Dutch, a verb-second language with a multifunctional clause-initial position, and English, which has a more rigid SVO structure (Los, 2009). It is against this background that this article aims to quantify and evaluate advanced Dutch EFL learners' use of clause-initial adverbials in a syntactically annotated longitudinal corpus of student writing, compared with a native reference corpus. Specifically, this paper zooms in on the use of adverbials of place and addition, which are both particularly relevant in this context as they are commonly used in Dutch to establish a link to the directly preceding context. In doing so, we hope to answer the following questions:

- 1) How do advanced Dutch EFL learners differ from native speakers in the frequency with which they use clause-initial place and addition adverbials and in the way they use these to provide a link to the preceding discourse?

- 2) Can a development be observed in advanced Dutch EFL learners' use of clause-initial place and addition adverbials in the direction of native writing?

Apart from providing further insight into the role of information-structural transfer in second language learning, the answers to these questions may have implications for language teaching at advanced stages of acquisition.

2. Comparative pragmatics of first position in Dutch and English

The first position has a key role in linking a clause to the preceding discourse and in providing a background against which the message in the remainder of the

clause is to be interpreted (Virtanen, 1992). In Dutch, a verb-second language, clause-initial position is both syntactically and information-structurally ‘multifunctional’, in the sense that it may be occupied by either a subject (1a), an object (1b) or an adverbial (1c) and, depending on context, these constituents may be either marked or unmarked (Los, 2009):

- 1) a. *Andy Cole heeft hier zijn eerste hattrick gescoord.*
 Andy Cole has here his first hat-trick scored.
- b. *Zijn eerste hattrick heeft Andy Cole hier gescoord.*
 His first hat-trick has Andy Cole here scored.
- c. *Hier heeft Andy Cole zijn eerste hattrick gescoord.*
 Here has Andy Cole his first hat-trick scored.

(Adapted from Hannay & Keizer, 1993)

In Dutch, clause-initial position commonly hosts what Los and Dreschler refer to as ‘local anchors’, adverbials which serve to link the sentence they occur in to the immediately preceding discourse (2012, p. 859). This tendency to link sentences together by means of (unmarked) clause-initial adverbials is shared by other V2 languages such as Norwegian (Hasselgård, 2009) and German (Kirkwood, 1969; Carroll *et al.*, 2000; Bohnacker & Rosén, 2008). English, on the other hand, has a more rigid SVO structure in which the use of clause-initial position is restricted and the subject has an important linking function (Carroll *et al.*, 2000; Los, 2009). Dreschler and Hebing (2010, p. 64), for example, find that 77% of English sentences start with a subject, while only 23% start with other first constituents

(against 54% subjects and 46% other first constituents for Dutch). It is clearly not ungrammatical in English for adverbials to occur in front of the subject, but, as Los (2009) argues, '[t]he very fact that this pre-subject position does not need to be filled, unlike the first position in a verb-second system, makes it likely that it has acquired a special, marked position' (26).

With 64 occurrences per 1,000 words (against 14 for initial position and 20 for medial position), in English 'final position is by far the most common position for adverbials' (Biber *et al.*, 1999, p. 772). This can be accounted for by the preference for this position that has been found for the large and varied class of circumstance adverbials, most of which are realised as prepositional phrases (Biber *et al.*, 1999, p. 772, p. 807). Interestingly, Biber *et al.* note that, especially in written registers, initial prepositional phrases may 'have a cohesive function, with the prepositional phrase using some information given in the previous discourse as the starting point for the next sentence' (1999, p. 809). This type of head-tail linking is exactly what we find in Dutch. However, differences between Dutch and English in the frequency with which such phrases are used and the contexts in which they commonly appear point to information-structural constraints associated with the use of prepositional phrases in clause-initial position in English sentences.

Translation manuals, such as Lemmens and Parr (1995, p. 92), for example, generally warn against using prepositional phrases at the start of an English sentence and suggest that the noun following the preposition in the Dutch adverbial (2a) should be used as a subject instead (see (2b) below), in order to avoid giving the adverbial, which was unmarked in Dutch, too much emphasis in its English translation, which is what happens in (2c):

38 - *Traces of Transfer*

2) a. *Met de geldautomaat heeft u altijd toegang*
With the cash dispenser have you always access

tot uw rekening.
to your bank account.

b. The cash dispenser gives you constant access to your bank account.

c. With the cash dispenser you always have access to your bank account.

(adapted from Lemmens & Parr, 1995)

Also consider the difference between the following sentences:

3) a. *In Nederland zijn de meeste scholen openbaar.*
In the Netherlands are the most schools public.

b. Education is usually public in the Netherlands.

c. In the Netherlands, education is usually public. (In Brazil, on the other hand, most middle-class children attend private schools.)

Compared with the clause-final adverbial in (3b), the clause-initial adverbial in (3c) is much more marked and is likely to be interpreted as contrastive. While Dutch circumstance adverbials, such as the instrument adverbial in (2a) and the place adverbial in (3a), are typically used in clause-initial position to function as

unmarked local anchors, in English the most likely reason to move them to clause-initial position would be to give them extra prominence or contrastive focus. It is the subject of the sentence which is the prime candidate for establishing an unmarked link to the preceding discourse.

Dutch students of English apparently lack awareness of these subtle cross-linguistic differences, which manifests itself in their typical overuse of pre-subject adverbials, particularly those that are meant to function as unmarked local anchors. Take, for example, the following sentences from a text by a Dutch student writer:

- 4) *Besides this*, I have also been active in my student society where I was responsible for the finance of two committees involved in the organization of our lustrum.
- 5) *Due to my internship* I realized that students are already capable of implementing their knowledge into business.
- 6) *Therefore*, I started the Student Consultancy Group which brings top students and companies together.

While the Dutch translations of the adverbials in italics could function as neutral discourse links, in English their position in front of the subject is more marked. A native speaker might have left out 'besides this' from sentence (4). 'Due to my internship' in (5) implies contrast due to its evocation of alternative sources of realization and might have been replaced by 'My internship made me realize', while 'Therefore' in (6) could have been replaced by a reverse wh-cleft: 'This is

why...'. In each case, it might be more natural to start the sentence with a subject rather than an adverbial.

3. Method

3.1 Corpus design

We evaluated the use of clause-initial place and addition adverbials in a longitudinal corpus of 899 student essays written by Dutch students of English language and culture between their first and third year at university. These were collected between 2008 and 2012 at Radboud University Nijmegen as part of the LONGDALE project, a European project aimed at compiling longitudinal corpora of texts written by advanced EFL learners with different language backgrounds (Meunier, 2015). As previous studies have provided clear evidence that the frequency with which different categories of adverbials are used in clause-initial position depends to a great extent on text type (cf. Biber *et al.*, 1999; Bohnacker & Rosén, 2008), we distinguished between two types of writing included in the Dutch component of LONGDALE: (1) 440 essays on various aspects of British or American literature and culture with a total word count of 481,956 and (2) 459 timed, argumentative in-class writing assignments on subjects such as 'the pros and cons of an obligatory stay abroad for BA students' or 'the need for conservation at Radboud University'. As the in-class assignments, which had to be completed within 30 minutes, are considerably shorter than the essays, the total word count for the in-class assignments adds up to 146,481. In order to allow an initial comparison to the level of academic writing that students of English language and culture are ultimately aiming to achieve, Philip Springer's VU Native Speaker Published Research Article Corpus (VUNSPRAC), consisting of 11 linguistics articles (79,121 words) and 11 literature articles (70,890 words), was used as a reference (Springer, 2012).

3.2 Procedure

Syntactic annotation was added to the corpora using the Stanford Parser (Klein & Manning, 2003), after which adverbial phrases occurring in pre-subject position in declarative main clauses were filtered out with Corpus Studio (Komen, 2012). Adverbials that were part of a quote were removed. This procedure resulted in a database of 8,774 clause-initial adverbials for the learner corpus (13.96 per 1,000 words) and 1,891 for the reference corpus (12.61 per 1,000 words). These adverbials were then categorised according to their semantic function and referential state (cf. chapter 1), to allow quantification of the use of place and addition adverbials in the different corpora, both overall and those used for local anchoring.

4. Results and discussion

4.1 Text types

The data show a clear distinction between the essays and the in-class assignments. While the essays typically use many place adverbials in pre-subject position to establish a link with the preceding context by means of references to the text that is being discussed, addition adverbials predominate in the in-class assignments. Both categories will be discussed in turn.

4.2 Place

While Dutch place adverbials commonly occur clause-initially to provide a link to the preceding context, previous research by Biber *et al.* has shown that in English only 5% of place adverbials occur in clause-initial position, another 5% in medial position, while an overwhelming 90% occur in clause-final position (Biber *et al.*,

1999, p. 802).³ This difference between Dutch and English is reflected in a steady decline in the use of clause-initial place adverbials in the LONGDALE corpus between year 1 and 3:

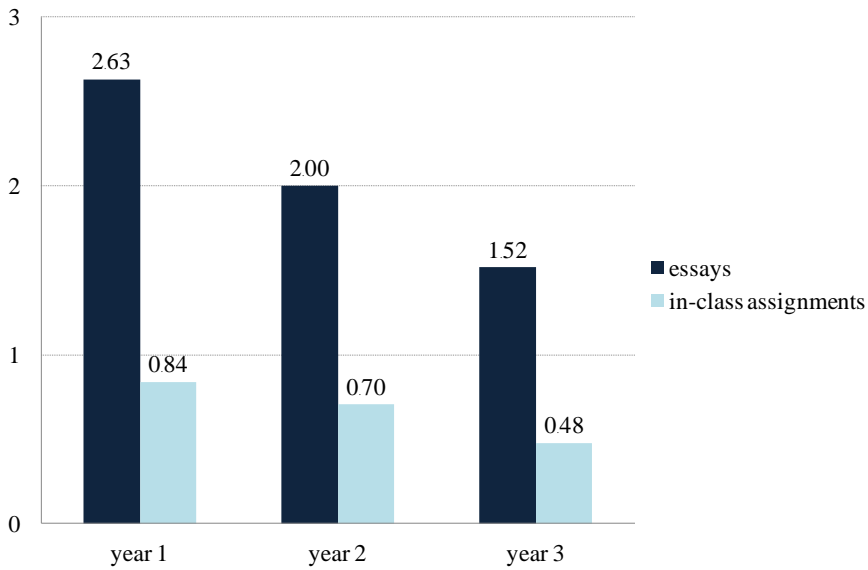


Figure 1. Clause-initial place adverbials per 1,000 words in LONGDALE

While a decline can be observed for both text types, they have widely different starting points. The essays start out at 2.63 place adverbials per 1,000 words in year 1, more than three times as many as for the in-class assignments, which start out at 0.84 per 1,000 words. This can be explained by the high number of references to the text that is being discussed in the literature essays, as exemplified by the following excerpts:

³ As the focus of the present study is on clause-initial adverbials, rather than the distribution of adverbials across initial, medial and final position, we cannot directly compare these results to our own data. The Longman Corpus used by Biber *et al.* therefore does not serve as reference material. However, it does serve to illustrate that English place adverbials have a clear preference for final position.

- 7) Besides celebrating their own identity, they used their works for promoting the spirit to keep going and to keep fighting for their rights. **In mother to Son**, Hughes uses the image of a mother telling her son to keep going and not to quit or give up.
(student RAD0801, year 1)

- 8) The proverb, “Alwey the nye slye maketh te ferre leeve to be looth”, speaks of a nearby sly person. **A couple of lines down**, Nicholas is described as nye and hende.
(student RAD0803, year 2)

- 9) He writes about things people would rather not read about; he writes about body and soul being always attached to each other, but writing about the body was a taboo at the time. **In section 11 in his Song of Myself** he writes about a widow desiring young men on the beach.
(student RAD0907, year 1)

The literature essays also include more place adverbials that are used for local anchoring, creating cohesion by linking back to an antecedent in the directly preceding context:

- 10) I'm talking about Hamlin Garland 's Under the Lion 's Paw, it is a great representation of Manifest Destiny where people settle somewhere to build their lives up again. **In this story** you see how hard it is for a farmer on his new piece of land to start over, that it is not all that easy, so farmers in the area help each other.
(student RAD1168, year 1)

11) The third crucial element of American romanticism are Short Stories. **In the short stories** the writers use the American landscapes and the typical American characters, which cut the connections with European history and made way to start their own new literature with adaptations to the environment of the country and the society.

(student RAD1182, year 1)

12) In 1855, the first edition of *Leaves of Grass* appeared, written by Walt Whitman, in which a poem called *Song of Myself* appeared. **In this poem** Whitman first describes that 'myself' in this poem is attached to everyone else: "and every atom belonging to me, as good belongs to you".

(student RAD1183, year 1)

Figure 2 visualises the decrease between year 1 and year 3 in the number of clause-initial place adverbials that function as local anchors. The essays go from 0.64 in year 1 to 0.37 in year 3, with a slight dip in year 2, but remain well above the level of the literature articles in the reference corpus, which use only 0.24 clause-initial place adverbials with an identity link to -1, i.e. the directly preceding sentence, per 1,000 words (Figure 3). The frequency of clause-initial place adverbials used in the in-class assignments, on the other hand, both overall and those that link back to -1, is considerably lower than the frequencies found in the reference corpus.

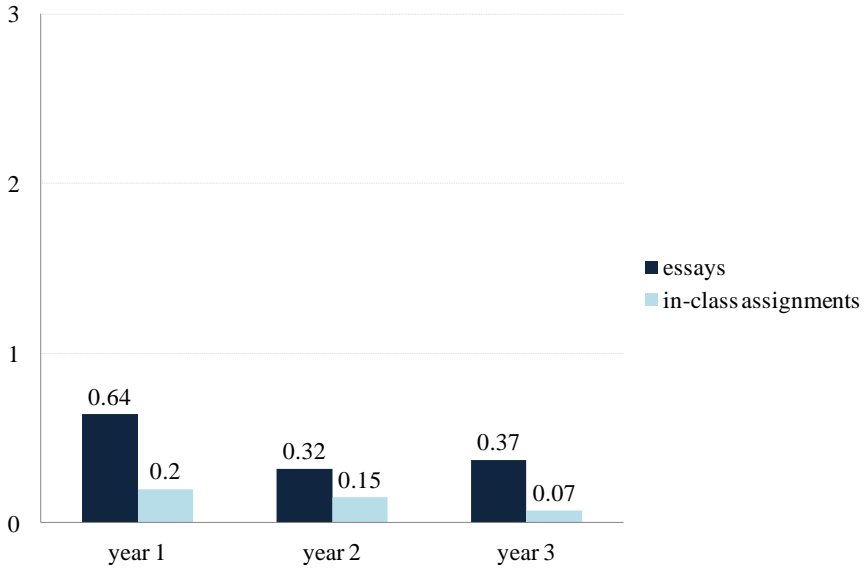


Figure 2. Clause-initial place adverbials with an identity link to -1 per 1,000 words in LONGDALE

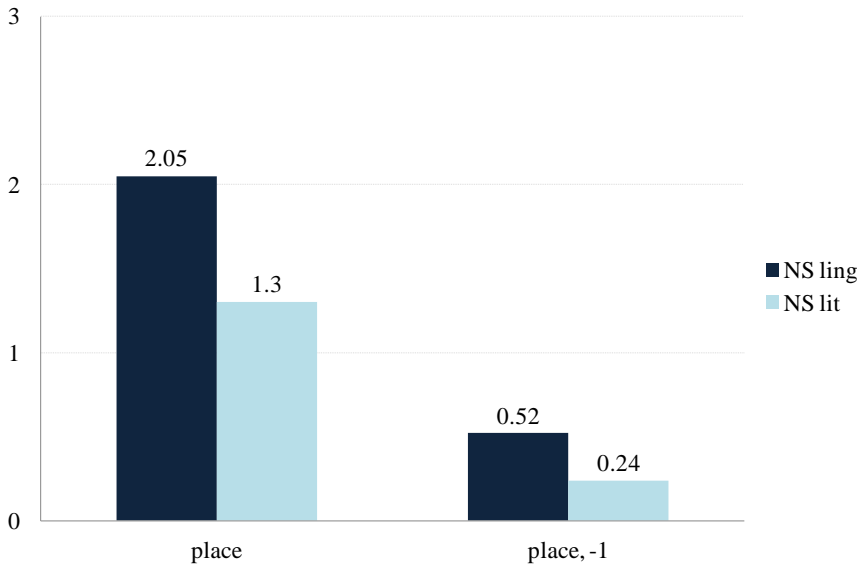


Figure 3. Clause-initial place adverbials and place adverbials with an identity link to -1 per 1,000 words in VUNSPRAC

Differences between the two types of texts included in the reference corpus also point towards a strong effect of genre. A closer look at the use of place adverbials in the linguistics and literature articles suggests that the literature articles are much less overtly structured than the linguistics articles. Most of the place adverbials in the linguistics articles are used for forward referencing (examples 13 and 14) or references to examples, tables etc. (example 15):

- 13) **In this section**, the traditional division drawn between internal and external explanations of language change is discussed.

(NS ling_01)

- 14) **In the following section**, discourse, narrative, and interpersonal relations are discussed [...].

(NS ling_02)

- 15) **In table 2**, raw frequencies are enriched by a significance test in the form of log-likelihood.

(NS ling_04)

4.3 Addition

Although the number of adverbials in clause-initial position in general might be expected to decline due to decreasing interference from Dutch, Figure 4 shows that the use of clause-initial adverbials that fall into the category of addition actually increases between year 1 and 3, only slightly for the essays, but more considerably for the in-class assignments, which start out at 1.43 clause-initial addition adverbials per 1,000 words and end up at 1.71 by year three. This might

be attributed to the fact that in English, too, addition adverbials, which fall into the class of linking adverbials, have a preference for initial position, especially in academic prose (Biber *et al.*, 1999, p. 891), and students are explicitly taught to structure their writing and to make sure they link their sentences together by means of cohesive devices.

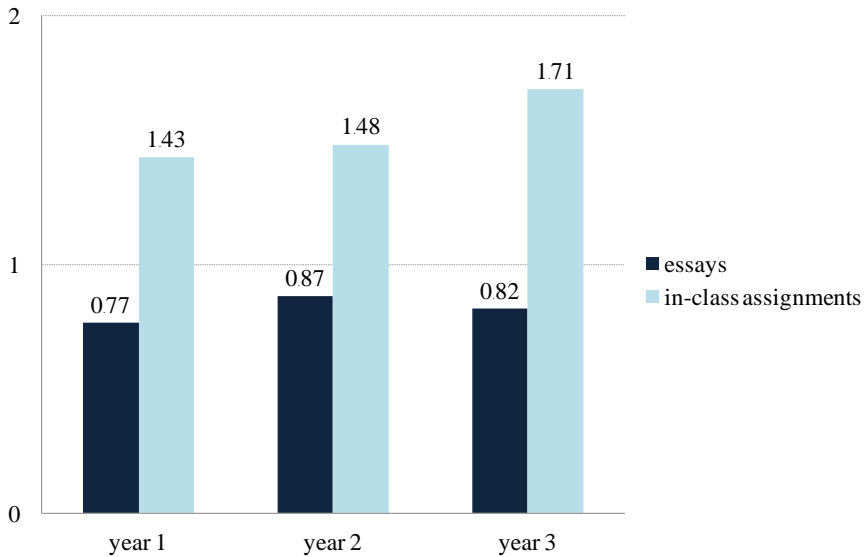


Figure 4. Clause-initial addition adverbials per 1,000 words in LONGDALE

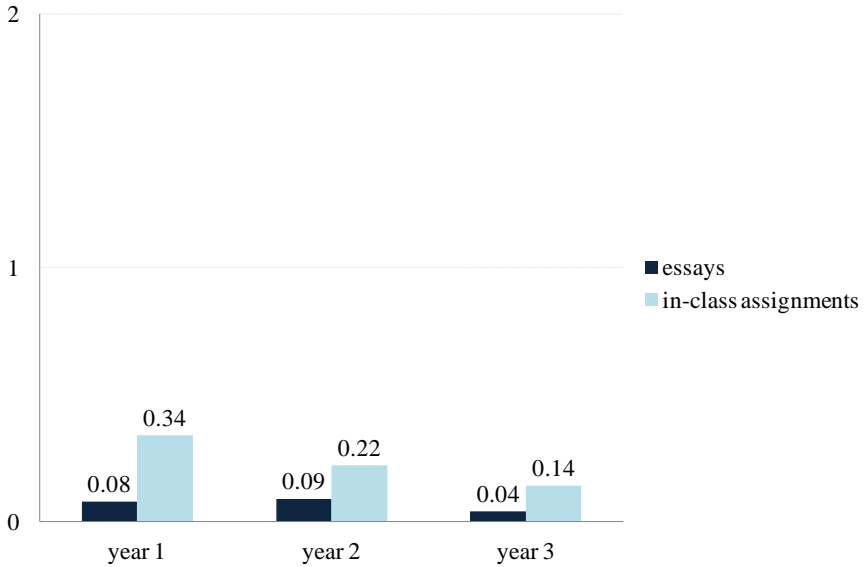


Figure 5. Clause-initial addition adverbials with an identity link to -1 per 1,000 words in LONGDALE

While there is an increase in the number of clause-initial adverbials in the category of addition overall, a different picture emerges from a sub-classification according to referential state. The learner corpus includes numerous examples of addition adverbials that are used for local anchoring, such as the following:

- 16) Women also were not allowed to write and convey their opinions. However, these women rebelled against this prison-like image. **In addition to this**, the fact that women were inferior to men, showed what America truly was.
(student RAD0927, year 1)

- 17) The stories of King Arthur have fascinated people for centuries. Until this day, they are thought very entertaining. But **besides that** they serve another important cause; they tell us about the people who wrote and were supposed to read it.
(student RAD0837, year 2)

Figure 5 shows that there is a steady decrease in the number of these addition adverbials which explicitly establish an identity link to the directly preceding context.

Interestingly, the number of addition adverbials in initial position in the reference corpus is in fact lower than the starting level for the category of addition adverbials in year 1 in the learner corpus:

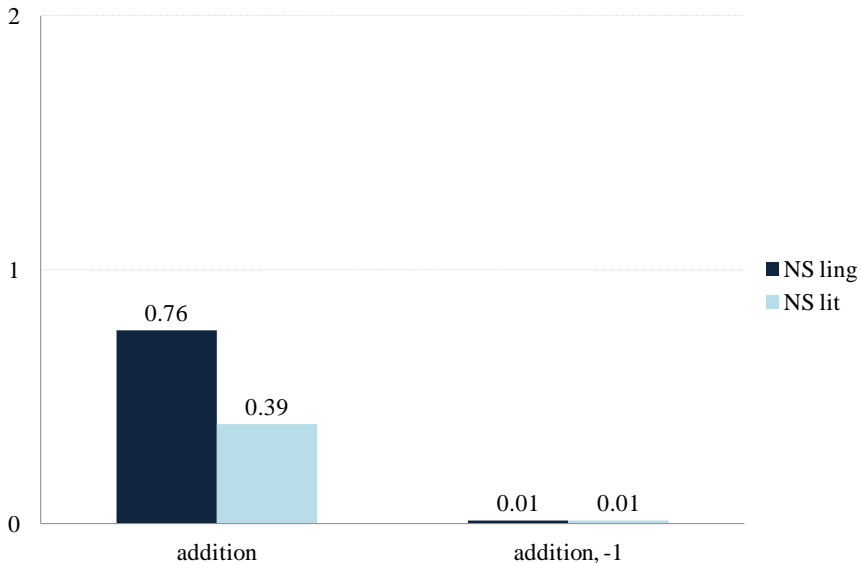


Figure 6. Clause-initial addition adverbials and addition adverbials with an identity link to -1 per 1,000 words in VUNSPRAC

Figure 6 shows that the native-speaker literature articles (NS lit) and the native-speaker linguistics articles (NS ling) use 0.39 and 0.76 addition adverbials per 1,000 words respectively, whereas the essays in the learner corpus (Figure 4), which are similar in genre to the NS literature articles and might therefore be expected to be comparable in style, start out at 0.77 and increase up to 0.82 by year three. The number of addition adverbials used in the in-class assignments is even higher, starting out at 1.43 and increasing up to 1.71 by year three. Although the reference corpus does not contain any writing that can be said to belong to the same genre as the timed, argumentative in-class writing assignments included in the learner corpus, it is clear that the number of initial addition adverbials in the in-class assignments is not just higher than that of the essays in the learner corpus but also higher than that of either category in the reference corpus.

What is even more striking is the fact that the category of ‘addition, -1’ is virtually non-existent in the reference corpus (Figure 6), which means that while there is a downward trend for the learners of English in the direction of native writing in this respect (see Figure 5), by year 3 there is still a gap between the number of addition adverbials with an identity link to -1 in the essays and in-class assignments in the learner corpus (0.04 and 0.14 per 1,000 words respectively) and the number of addition adverbials in both the linguistics and the literature articles in the reference corpus (0.01 per 1,000 words for both groups).

5. Conclusion

The case study presented here shows that differences between Dutch and English in the information status of clause-initial position appear to lead advanced Dutch EFL learners to overuse clause-initial place or addition adverbials, which in Dutch are commonly used to provide a link to the preceding discourse. However, there are interesting differences between these two categories. Place adverbials, like

most circumstance adverbials, only rarely occur in pre-subject position in English sentences, of which students seem to become increasingly aware between their first and third year at university. The number of place adverbials that refers back to the directly preceding context declines at a similar pace. Addition adverbials, on the other hand, belong to the class of linking adverbials, which in fact have a preference for initial position in English as well. As a result, students seem to be less aware of the markedness of the frequent use of addition adverbials in clause-initial position in their L2, judging from the increase in this category between year 1 and year 3. Interestingly, a development in the direction of native writing can only be observed when addition adverbials are subcategorised according to referential state, which brings to light a steady decline in the number of addition adverbials that have an identity relation with an antecedent in the directly preceding context. Still, it is in this sub-category that the biggest contrast between native writing and the writing of advanced Dutch EFL writers can be observed, even by the end of year 3, as addition adverbials in the reference corpus tend to achieve cohesion by means of implicit reference (e.g. *in addition, moreover*), while the Dutch learners continue to establish explicit links (e.g. *in addition to this, on top of that*) that serve to anchor their sentences in the directly preceding discourse.

These results are of course particularly relevant for those who, like our students of English Language and Culture, are not just EFL users but future EFL professionals (De Haan & Van der Haagen, 2012). Many researchers have observed that exposure to the target language alone is not sufficient to learn to recognise subtle cross-linguistic differences (Norris & Ortega, 2000; Hinkel, 2003). If non-native writers wish to reach beyond grammatical correctness and acquire a near-native level of proficiency, they will need teaching materials focusing on those areas which are still vulnerable to L1 influence even at advanced stages of acquisition. At present, most textbooks are not informed by corpus research. The

few exceptions to the rule are based on native-speaker corpora and not specifically aimed at non-native writers (Gilquin *et al.*, 2007). Further corpus research into transfer of language-specific principles of information structure could therefore serve as an empirical basis for the development of L1-specific EFL teaching materials.

3. Common framework, local context, local anchors: How information-structural transfer can help to distinguish within *CEFR C2*⁴

Abstract

This paper addresses the problem of defining language development at advanced stages of L2 acquisition. It is commonly assumed that the differentiation between and within C1 and C2 is mostly determined by the learners' educational and cognitive development. While it is true that it is probably impossible to distinguish the higher levels on the basis of linguistic development only, we argue that it does continue to play a subtle role, particularly in the form of L1 information-structural transfer. We believe, therefore, that L1 background should be taken into account in defining language development at higher levels of L2 acquisition. Following up on the study presented in chapter 2, we present an analysis of a sub-corpus of the Dutch component of LONGDALE (Meunier, 2015), which was expected to reveal traces of V2 to SVO information-structural transfer in the form of pre-subject adverbials that function as 'local anchors' (Los, 2012), providing a link to an antecedent in the directly preceding discourse. This sub-corpus was aligned with the *Common European Framework of Reference for Languages (CEFR)* by means of an Oxford Online Placement Test (OOPT, Purpura, n.d.) which students took in the week of the LONGDALE writing assignment. The statistical analysis of our data supports the idea that Dutch learners of English use more local anchors than native speakers who are comparable in age and academic background. Also,

⁴ This is a revised version of: Van Vuuren, S. & De Vries, R. (forthcoming). Common Framework, Local Context, Local Anchors: how V2 to SVO information-structural transfer can help to distinguish within *CEFR C2*. In B. Los & P. De Haan (Eds.), *Word Order Change in Acquisition and Language Contact*. Amsterdam: John Benjamins.

students' OOPT scores were found to be inversely correlated with their use of local anchors, suggesting that information-structural transfer is a feature that can be used as an indicator of language development at higher levels of L2 acquisition. Finally, a suggestion is made for an L1-specific local descriptor of coherence and cohesion to capture this feature of Dutch (future) English language professionals' proficiency.

1. Introduction

1.1 The Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)

The *CEFR* was officially published in 2001 and was hailed by some, or many, as a watershed in language learning and teaching. Capitalising on the communicative language teaching revolution that had shaken the teaching world to its foundations in the 1970s and 80s, the *CEFR* more or less authoritatively moved the focus away from the much-maligned knowledge about a language to the things a learner should be able to do in that language; the achievement of communication was officially given pride of place over the achievement of accuracy.

The *CEFR* was a long time in the making. The first awareness of the need of some sort of standard developed in a post-war Europe that had become convinced that European languages were not only worth preserving as a rich linguistic and cultural heritage, but also an essential means of preventing further conflicts and promoting understanding across (linguistic) borders. The final push towards the *CEFR* was made during a conference in Rüschiikon (Switzerland) in November 1991. Its 229-page report concluded there was a need for 'a comprehensive, coherent and transparent framework for the description of language proficiency' (Council of Europe, 1992).

Although the *CEFR* may be seen as a watershed in language teaching, it has not been without its critics. It has been called an ‘unsafe framework that is failing language learners across Europe’ (Fulcher, 2004) and has been accused of being built on ‘shaky ground’, combining mixed ‘what’ and ‘how well’ descriptors (Hulstijn, 2007). It has been criticised for not being theoretical, for being too intuitive (Hulstijn, 2007), for ignoring findings from SLA research and for not catering for learning diversity along the proficiency continuum (Hulstijn qtd. in Figueras, 2007). It has also been pointed out that the scales and descriptors cannot be used for very young learners, nor for English for Specific Purposes (ESP) and Content and Language Integrated Learning (CLIL) settings without supplementary descriptors (Little, 2007). It has been called both vague, repetitive, too abstract, lacking in detail and unclear (Martyniuk & Noijons, 2007) as well as restrictive (McNamara, 2007). Despite claims by stakeholders, such as examination providers and publishers, that their products are aligned with the *CEFR*, no theory of comprehension has been formulated that can actually measure the ‘mental operations’ that the L2 speaker has to demonstrate at any particular *CEFR* level (Alderson, 2007). The Dutch *CEFR* Construct Project concluded that the *CEFR* descriptors in their current form cannot be used for the construction of language tests (Alderson *et al.*, 2006).

For the purpose of this paper, there are two main issues that we would like to address: the underspecification of C1 and C2 and the minimal role for the learners’ L1. As has been noted before, the *CEFR* is not very good at distinguishing at the higher levels of proficiency. The total number of scales is 52, but 10 descriptors at C1 and as many as 25 at C2 are missing from those scales or described as equal to a lower scale (cf. Springer, 2012). Secondly, the role of the learner’s L1 is only mentioned in four of the 52 scales, and then only for one descriptor at a lower level:

Compensating

B1: Can foreignise a mother tongue word (...)

Grammatical accuracy

B1+: (...) generally good control though with noticeable mother tongue influence (...)

Orthographic control

B2: (...) Spelling and punctuation are reasonably accurate but may show signs of mother tongue influence

Phonological control

A2: Pronunciation is generally clear enough to be understood despite a noticeable foreign accent (...)

(Council of Europe, 2001)

Of course, a ‘common’ European framework is by necessity non-language specific, both for L1 and L2. For this reason, the *CEFR* specifically invites users to adapt the framework to their local context. Some progress is being made in this respect with the development of so-called Reference Level Descriptions for a variety of L2s (Hawkins & Filipovic, 2012). Apart from the L2 itself, most language teachers will agree that a learner’s L1 also plays a role in the development of that L2, most clearly seen at the lower proficiency levels, but also, more subtly, at higher levels. This would seem to suggest that adapting the framework to the local context of the learner’s L1 could help in addressing some of the problems described above.

1.2 The *CEFR* in the Dutch educational system

In the Netherlands, both secondary and higher education formulate their exit levels in *CEFR* terms. At VWO-level (roughly equivalent to British A-levels),

secondary school exams for the major foreign languages are aligned in the following way:

Table 1. Exit levels for foreign languages in pre-university education (adapted from ERK, n.d.)

	German	English	French	Spanish
Listening	B2	B2	B2	B2
Spoken production	B2	B2	B1+	B1+
Spoken interaction	B2	B2	B1+	B1+
Writing	B1	B2	B1	B1
Reading	B1 (20%) ⁵	B2 (85%)	B1 (60%)	B1 (60%)
	B2 (75%)	C1 (15%)	B2 (40%)	B2 (40%)
	C1 (5%)			

This means that Departments of English in Dutch universities expect their students to come in at B2. They also specify their exit levels:

Table 2. Exit levels of English departments in the Netherlands

University	Written	Spoken interaction	Spoken production	Reading	Listening
Amsterdam UvA	C1	C1	C1	C2	C1
Amsterdam VU	C1	?	?	?	?
Groningen	C1	C1	C1	C1	C1
Leiden	C1	C1	C1	C2	C1
Nijmegen	C2	C2	C2	C2	C2
Utrecht	C1	C1	C1	C2	C1

⁵ Percentages refer to approximate proportion of exam questions at each level.

It is interesting that only Radboud University, Nijmegen, assumes that its students graduate at C2 across the board. All sister departments see their students leave at C1, except for reading, which some departments believe to be at C2.

Until 2013 the English department at Radboud University had just taken both the B2 entry level and the C2 exit level as ‘read’. In September 2013 it was decided to subject all first-year students to the most widely-used, *CEFR*-aligned placement tool, the OOPT. When the results were returned, only a handful of students turned out to be at the expected B2, with roughly 40% at C2, the actual projected exit level.

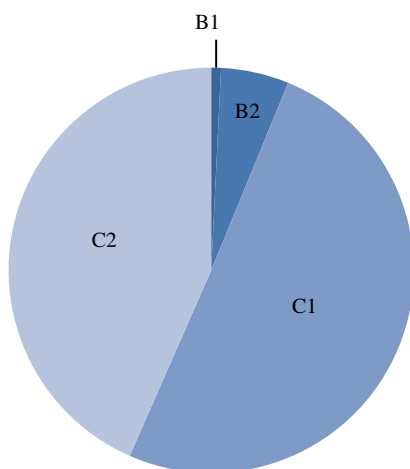


Figure 1. *CEFR* entry levels first-year students of English at RU (September 2013)

It was considered inconceivable that so many first-year students should be at the exit level of third-year students, who in the course of their time in the department would have taken a suite of rigorous written and oral proficiency courses, as well as having been constantly exposed to English during lectures and seminars. While this raises obvious questions about test validity, it is also likely to reflect the

underspecification of the higher *CEFR* levels, C2 in particular, which represents such a vast expanse of language, language skills and cognitive skills, that it is impossible to say where exactly, or even roughly, in this expanse a particular student is. The following inverted triangle, which visualises the development of business English skills, illustrates this point particularly well. It demonstrates the vastness of the area, as well as the other skills involved at the highest level.

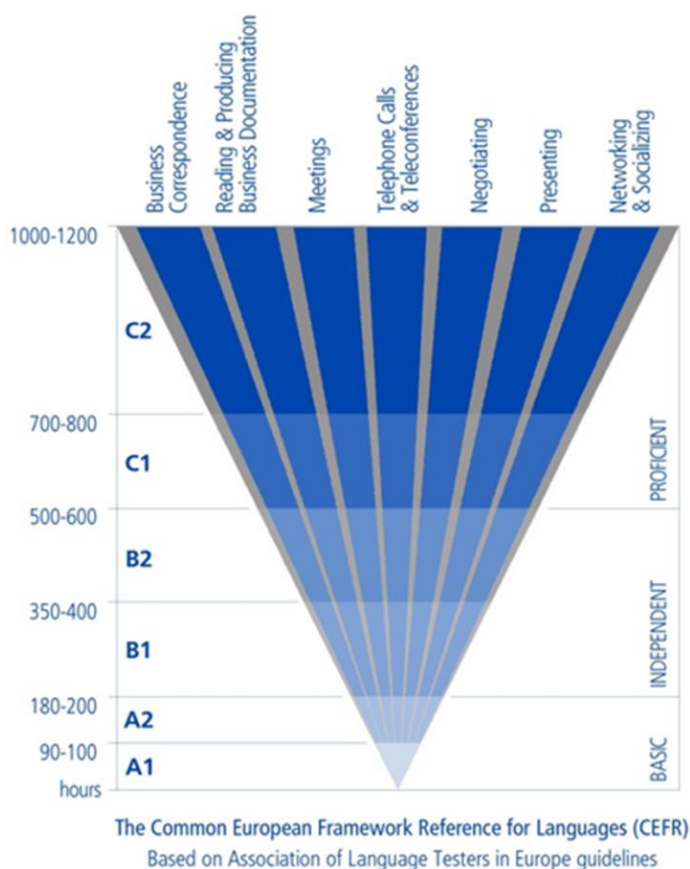


Figure 2. Development of business English skills A1-C2 (Target Training, n.d.)

The vastness of the area occupied by C2, in combination with the fact that many descriptors are missing at these higher levels, clearly highlights the need for research to establish whether there are L1 features, transfer from Dutch in this case, that might allow us to tentatively define different stages of development within C2. At the same time, we also need to recognise that different language users have different needs. Students of English are educated to become English language professionals, such as teachers, language coaches and editors, rather than straightforward language users (De Haan & Van der Haagen, 2012). Their needs are therefore different from those of most language users. The *CEFR* itself is aware of the need to adapt the framework to the needs of different groups of learners: '[establishing] broad- or narrow-band levels in accordance with their need to make finer or coarser distinctions among a population of learners' is exactly what the *CEFR* expects its users to do (Council of Europe, 2011, p. vi).

1.3 L1 transfer at higher levels of acquisition

It is probably true that it is impossible to distinguish the higher levels on the basis of linguistic development only. We nevertheless argue that linguistic development does continue to play a subtle role even at higher levels of acquisition, particularly in the form of information-structural transfer from learners' L1. The results presented in the previous chapter suggest that Dutch students of English are likely to show traces of V2 to SVO information-structural transfer in their use of pre-subject adverbials, more specifically in their preference for linking sentences by means of pre-subject adverbials that function as 'local anchors' (Los, 2012), providing a link to an antecedent in the directly preceding discourse, as illustrated by the following example by a first-year student writer:

- 1) To make sure that first-year students will be successful in college they need to take certain steps. To start off with, every student needs to have a good day-planner in which they can note down the appointments with workgroups, due dates of assignments, tests, reading assignments etc. **Via this day-planner**, students can make a clear and concise schedule from which they can work on.

(student RAD0019)

The adverbial, ‘via this day-planner’, occurs in pre-subject, clause-initial position to provide a link to the preceding sentence, where the day-planner was introduced. This type of cohesive strategy is a typical feature of Dutch, which has a multifunctional preverbal position that can accommodate both marked and unmarked subjects, objects and adverbials. While discourse linking by means of local anchors is syntactically possible in English, unmarked links are typically realised by the grammatical subject of the sentence and pre-subject adverbials, such as in the example above, tend to be much more prominent than their unmarked Dutch equivalents (cf. Los, 2012; Verheijen *et al.*, 2013).

The study reported on in chapter 2 looked at four cohorts of students who started between 2008 and 2011, focusing on both place adverbials, as an example of a circumstance adverbial, and addition adverbials, as an example of a linking adverbial. Both categories were further subdivided into those adverbials that referred back to an antecedent in the directly preceding discourse (i.e. that functioned as local anchors) and all other adverbials. The learner corpus was found to contain more place and addition adverbials than the VU Native Speaker Published Research Article Corpus (Springer, 2012). The use of place adverbials turned out to decline between year 1 and 3. This applied to place adverbials overall and local anchors alike. Addition adverbials, on the other hand, were shown to

increase between year 1 and year 3, in spite of the fact that even in year 1 learners already used more addition adverbials than native speakers. A development in the direction of native writing was only found for the sub-category of local anchors: the use of phrases like ‘in addition to this’ clearly decreased between year 1 and 3. The different patterns observed for place and addition adverbials might be explained by the fact that in English place adverbials have a preference for final position, which students might gradually become aware of, while for linking adverbials ‘initial position can [...] be considered the unmarked position’ (Biber *et al.*, 1999).

These findings suggest that advanced Dutch EFL writers betray transfer from their L1 in the way they use local anchors for discourse linking and that a decrease in the use of local anchors can be used as a marker of language development. However, as students’ ‘advancedness’ was determined on the basis of institutional status only, it is not clear at what stage students become more native-like in this respect and whether or not this feature of learner language can be used to define language development at the higher end of the *CEFR*. The present study, therefore, has three objectives: (1) to establish whether Dutch learners’ tendency to overuse local anchors, as described by previous studies, is statistically significant, (2) to investigate whether this tendency diminishes with increasing proficiency, operationalised by students’ OOPT scores, and (3) to see whether the use of local anchors can function as an objective indicator of language development within *CEFR* level C2.

2. Materials and Methods

2.1 Corpus design

We evaluated the use of pre-subject adverbials in a 48,608 word sub-corpus consisting of 118 untimed writing assignments on ‘how to be successful in

college/university' produced by Dutch students of English language and culture at the beginning of their first year. The texts were collected at Radboud University Nijmegen as part of the LONGDALE project (Meunier, 2015) in September 2013, just after students had taken an OOPT.

Students' writing assignments were compared with a 60,739 word reference corpus consisting of 114 A-level argumentative essays included in LOCNESS (2005), as these are similar in their level of formality and their authors can be expected to be comparable in academic background and maturity.

2.2 Procedure

The procedure that was used is an elaboration on the procedure described in the previous chapter and is depicted schematically in Figure 3. Both corpora were parsed using the Stanford Parser (Klein & Manning, 2003), after which adverbial phrases occurring in pre-subject position in declarative main clauses were identified with Corpus Studio (Komen, 2012). This resulted in a database of 579 pre-subject adverbials for the LONGDALE sub-corpus and 723 for the reference corpus, both of which were then annotated in Cesax (Komen, 2012) by assigning a semantic label (e.g. 'instrument', 'addition', 'place') to each adverbial and adding information about its referential state (cf. chapter 1).

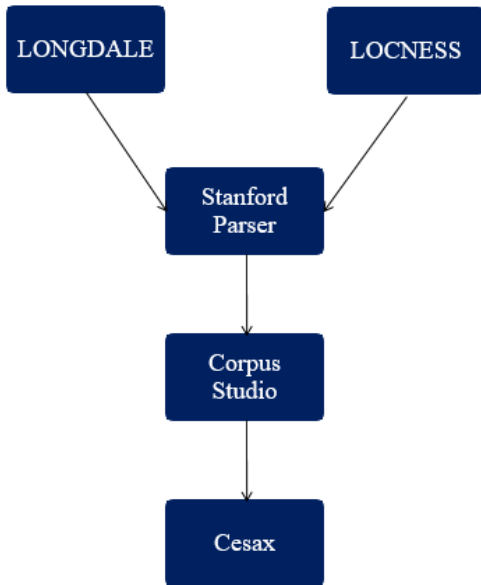


Figure 3. Schematic overview of procedure

After removal of quotes, the database contained 499 pre-subject adverbials for the LONGDALE sub-corpus and 558 for the LOCNESS A-levels. Biber *et al.* (1999) group the many different semantic categories of adverbials they distinguish into three main classes of adverbials – circumstance adverbials, stance adverbials and linking adverbials (cf. chapter 1, Table 3). For the current study, we considered pre-subject adverbials overall and each of the three main adverbial classes separately, as each class has different properties and preferences for use in initial, medial or final position. We also looked at those adverbials – regardless of adverbial class – that are used for local anchoring, i.e. that include an identity link to an antecedent in the directly preceding context.

Students who did not complete the OOPT were excluded from further analysis, so that our final database included 461 pre-subject adverbials in 109 texts for the LONGDALE sub-corpus and 558 pre-subject adverbials in 111 texts for the LOCNESS A-levels. The use of each of the three adverbial classes was then

quantified with another Corpus Studio query and results were prepared for analysis in SPSS by means of Cesax (Komen, 2012).

2.3 Normalisation

Like Verheijen *et al.* (2013), the study reported on in chapter 2 normalised raw frequencies per x number of words. However, we have since come to believe that this makes it difficult to disentangle differences in the number of pre-subject adverbials and differences in sentence length, which are likely to exist between learners and native speakers as well as between learners in successive years. In theory, an increase in sentence length between year 1 and year 3 could create the impression that use of pre-subject adverbials is going down. As each main clause offers one slot which the writer can choose to fill with a pre-subject adverbial or not, we decided to opt for normalisation per 100 main clauses rather than per 1,000 words, in order to control for differences both in text length and in sentence length between our learner corpus and reference corpus.

2.4 Statistical analysis

Both Verheijen *et al.* (2013) and the study reported on in the previous chapter considered the use of pre-subject adverbials in LONGDALE in a longitudinal perspective, making statistical analysis difficult due to the (possibly selective) loss of participants between year 1 and 3. The current study therefore also serves to verify whether the differences between advanced learners and native speakers found in those studies are statistically significant, before considering the correlation between language proficiency (operationalised as OOPT score) and use of pre-subject adverbials. Our data were not normally distributed, especially in the case of local anchors, which, in some texts, are not represented at all. For this reason, we used a non-parametric ANOVA – an independent samples Mann-

Whitney U-test – to compare the learner and reference corpora. We then used Spearman correlation for non-parametric data to investigate the relationship between students’ OOPT scores and the number of pre-subject adverbials used per 100 main clauses.

3. Results

3.1 Native speakers and Dutch students of English compared

The results show that the mean number of pre-subject adverbials per 100 main clauses used by the Dutch learners ($M = 5.97$, $SD = 3.42$) is slightly lower than that of the native speakers ($M = 6.64$, $SD = 3.94$), although this effect is not significant ($p = .184$):

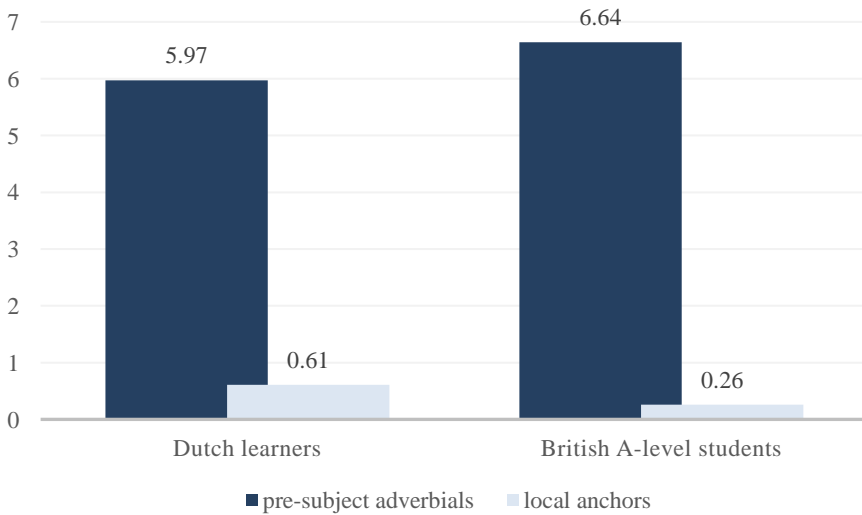


Figure 4. Pre-subject adverbials overall and local anchors per 100 main clauses

Breaking down this relatively lower mean number of pre-subject adverbials into the three main classes of circumstance, stance and linking adverbials (Figure 5) shows that it is the stance adverbials in particular which the learners use fewer of

($M = 0.73$, $SD = 1.10$) in comparison with the native speakers ($M = 1.20$, $SD = 1.31$). This lower use of stance adverbials is significant at $p = .004$. The learners' relatively lower use of circumstance adverbials ($M = 2.14$, $SD = 2.02$ for the learners, against $M = 2.73$, $SD = 2.24$ for the native speakers) is not significant ($p = .060$), nor is their relatively higher use of linking adverbials ($M = 3.10$, $SD = 2.75$ for the learners, against $M = 2.71$, $SD = 3.12$ for the native speakers, $p = .132$). Most importantly, as visualised in Figure 4, our results confirm the idea that the learners use significantly more local anchors ($p = .005$).

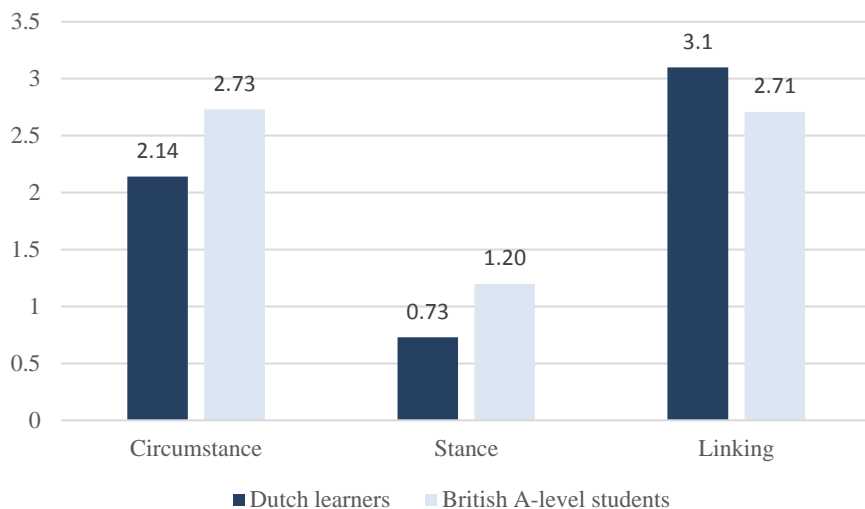


Figure 5. Pre-subject adverbials broken down according to adverbial class

The mean use of local anchors of the Dutch learners ($M = 0.61$, $SD = 0.97$) is considerably higher than that of the native speakers ($M = 0.26$, $SD = 0.58$). The local anchors in our learner corpus are either circumstance or linking adverbials. Within these two adverbial classes, the categories of place adverbials (examples (2) and (3)) and addition adverbials (examples (4) and (5)) feature most heavily, accounting for 31% and 24% of all local anchors respectively. Other adverbial

categories that are used as local anchors include time, domain, means, instrument, purpose, contrast, cause/reason and concession. As illustrated by examples (4) and (5), the pronominal adverbs (equivalent in this case to the Dutch *daarnaast* and *bovendien*) that were identified by Verheijen *et al.* (2013) as a characteristic of advanced Dutch EFL writing also feature heavily in the category of local anchors.

- 2) Another way to get prepared is to read through the course guides. **In these guides** everything is written that students need to know about their courses, such as literature needed, contents of the course and the kinds of exams it will have.

(student RAD1355)

- 3) Also you should attend all lectures and seminars, even though you are not obligated to go to them. **In these lectures and seminars** the teachers will tell you a lot about the subject and they can explain the things you did not understand before.

(student RAD1399)

- 4) If students understand the subject-matters, they will easily pass their exams, which is of course really important. **Besides that**, students could join different associations to meet new people.

(student RAD0030)

- 5) During that week first-years start building up a network of contacts, for they become acquainted with their classmates as well as with second or third year students, who are eager to give them information and advice about their study. **Above that**, first-years get a good impression of student parties.
(student RAD1370)

Apart from the fact that the native speakers use fewer local anchors, these local anchors are also restricted to the class of circumstance adverbials:

- 6) This is the system of direct constituencies. **Within these constituencies** the candidate with the largest number of votes is elected.
(alevels2)
- 7) The reasons these discussions have occurred is because of the serious injuries which have occurred in fights such as the tragic Benn-McLellan, and Eubank-Watson, matches. **In both of these clashes**, the result of the match has been a brain injury to one of the boxers.
(alevels4)
- 8) This will enable scientists to know precisely what each gene does. **With this information** certain genes can be searched for in the foetus and if detected the genes can be replaced with normal genes instead of the diseased ones.
(alevels8)

In our reference corpus linking adverbials are not used for local anchoring at all.

3.2 OOPT as a predictor of adverbial use

Figure 6 shows mean use of pre-subject adverbials in general, and local anchors in particular, for *CEFR* levels C1 and C2 (based on students' OOPT scores). The other *CEFR* levels were not included because, out of a total of 109 students, only one was shown to be at B1 level, while there were just four at B2. We therefore do not think that our findings can be generalised beyond the levels of C1 and C2.

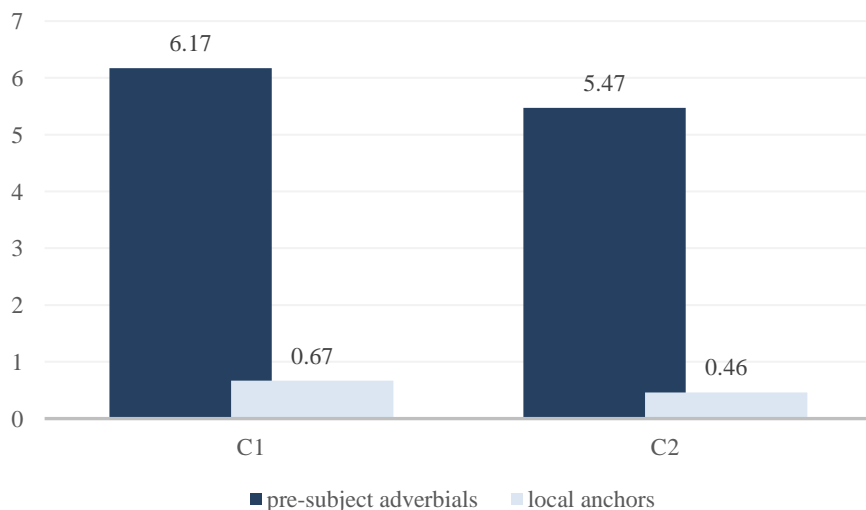


Figure 6. Use of pre-subject adverbials and local anchors at *CEFR* C1 and C2

With a mean use of 5.47 per 100 main clauses at C2 level against 6.17 at C1, more proficient learners appear to use slightly fewer pre-subject adverbials overall than less proficient learners, although there is no significant correlation between OOPT score ($M = 98.48$, $SD = 9.71$) and the total number of pre-subject adverbials used per 100 main clauses ($\rho(107) = -.168$, $p = .082$, two-tailed). There is, however, a weak but significant inverse correlation between OOPT score and the number of local anchors used per 100 main clauses ($\rho(107) = -.202$, $p = .036$, two-tailed),

meaning that more proficient EFL learners were slightly more native-like in their use of local anchors. Whereas learners at C1 level used an average of 0.67 local anchors per 100 main clauses, at C2 level this was 0.46 per 100 main clauses, which is closer to the native speaker mean of 0.26. There was no significant correlation between OOPT score and the separate adverbial classes of circumstance ($\rho(107) = -.004, p = .969$, two-tailed), stance ($\rho(107) = -.030, p = .756$, two-tailed), and linking ($\rho(107) = -.136, p = .157$, two-tailed).

4. Discussion

The purpose of this paper was to illustrate the use of *CEFR*-aligned learner corpora in defining language development at higher levels of L2 acquisition. To this end we needed (1) to confirm that Dutch learners of English do indeed use significantly more local anchors than native speakers, (2) to investigate whether OOPT score is correlated with use of local anchors, and (3) to see whether this feature could serve as an objective indicator for language development.

Our results show that there is no significant difference between the essays included in the Dutch component of LONGDALE and the A-level argumentative essays in LOGNESS in the number of pre-subject adverbials used per 100 main clauses. However, the Dutch learners are set apart from the native speakers on two counts. Firstly, the number of local anchors used per 100 main clauses differs significantly between both groups. Secondly, when native speakers do use adverbials for local anchoring, they invariably use circumstance adverbials, whereas the learner corpus also includes many linking adverbials that are used to explicitly refer back to an antecedent in the preceding discourse, phrases such as 'in addition to this'. This can probably be explained by the prevalence of their Dutch equivalents, pronominal adverbs such as *daarnaast* (next to that) and *bovendien* (on top of that).

A feature of the writing of Dutch learners of English which perhaps does not stand out as much – and which we did not expect on the basis of previous research – is that they use significantly fewer stance adverbials, adverbials like ‘unfortunately’, ‘naturally’, ‘hopefully’, or ‘basically’, which ‘convey speakers’ comments on what they are saying (the content of the message) or how they are saying it (the style)’ (Biber *et al.*, 1999). As we are looking at A-levels and beginning university students, this cannot be attributed to differences in academic maturity, nor is it likely to be a simple matter of proficiency as the adverbs that are part of the class of stance adverbials are not necessarily more linguistically complex than other adverbs. Instead, it might be due to the fact that, unlike linking adverbials and, to a certain extent, circumstance adverbials, stance adverbials do not serve to increase cohesion in a text and might for that reason be underemphasised in the writing curriculum.

Looking at students’ proficiency levels, operationalised as OOPT score, as a predictor of adverbial use, it appears that students at the higher end of the scale use fewer local anchors than students with lower proficiency scores. This represents a trend in the direction of native writing. Admittedly, the present study cannot be said to be longitudinal, as it does not chart our students’ development but looks at one collection of texts produced by one cohort of students who are at different levels of proficiency at a particular moment in time. For this reason, we plan to do a follow-up study in which we trace the same cohort’s writing development until the end of their third year, when they will take another OOPT and submit a final writing assignment. It will be interesting to see if (and when) students do start to adopt more stance adverbials at a later stage and whether the decrease in the use of local anchors persists.

Nonetheless, this cross-sectional view of information-structural transfer at different levels of language proficiency does suggest that the use of local anchors

can be used as an objective indicator for language development. We would therefore suggest that it would be useful to formulate an L1-specific descriptor in the form of a context-specific supplementary descriptor attached to the coherence and cohesion descriptor at *CEFR C2*:

Table 3. Supplementary coherence and cohesion descriptor.

PRAGMATIC COMPETENCE – DISCOURSE COMPETENCE		
COHERENCE AND COHESION		
C2	<i>CEFR</i> :	Can create coherent and cohesive text making full and appropriate use of a variety of organisational patterns and a wide range of cohesive devices.
C2	Dutch English language professionals:	Can create coherent and cohesive text without infelicitous use of local anchors.
		<ul style="list-style-type: none"> • Avoids linking adverbials with an explicit reference to an antecedent in the preceding discourse (i.e. linking adverbials that function as local anchors). • Only uses local anchors in a context where a marked discourse link is appropriate. • Uses the grammatical subject of the sentence as a default option for unmarked discourse linking.

In line with the *CEFR*'s invitation 'to make finer or coarser distinctions among a population of learners', this supplementary descriptor sets apart (future) Dutch English language professionals from other language users. We would ideally like to develop this descriptor into a tool allowing us to differentiate stages of language development within this group of learners, for example informing our

expectations of first- as opposed to third-year students of English at C2. We expect to be closer to an answer when the current cohort has completed its third year.

5. Conclusion

Our 2013 cohort study seems to suggest that taking L1 transfer into consideration may contribute to further specification of the higher levels of the *CEFR*. We demonstrated that the use of local anchors can function as an indicator of language development in our local context. More proficient users turn out to be more native-like in the sense that they use fewer local anchors. It is not beyond the realm of imagination to assume that other (information-structural) features may be employed to further specify the higher levels of the *CEFR*. We would therefore recommend the use of longitudinal *CEFR*-aligned corpora to identify other features of L1 transfer that might eventually lead to the identification of narrower bands in what is now a sea of C2.

4. Dutch learner English in close-up: A Bayesian corpus analysis of pre-subject adverbials in advanced Dutch EFL writing⁶

Abstract

This study presents a longitudinal Bayesian analysis of pre-subject adverbial (PSA) use in a 571,174-word corpus of Dutch learner English. Overall, learners use more circumstance and linking adverbials, but fewer stance adverbials than both novice and expert native writers. Learners also use more ‘local anchors’, adverbials that serve a cohesive purpose by linking back to an antecedent in the directly preceding discourse. Interestingly, untimed essays are more native-like than timed essays in their use of each of the adverbial types considered. Although learners generally develop in the direction of native writing, the use of linking adverbials in learners’ literature essays develops in the opposite direction. L1 transfer may account for more frequent use of some categories of adverbials, particularly local anchors, but the widening gap between learners and native speakers (NSs) in the use of linking adverbials suggests that the potential role of textbook and curriculum design warrants further research.

Keywords: information structure, discourse linking, EFL, advanced learners, L1 transfer

⁶ This is a revised version of: Van Vuuren, S. & Laskin, L. (2017). Dutch learner English in close-up: use of pre-subject adverbials as a defining feature of advanced Dutch EFL writing. *International Journal of Learner Corpus Research*, 3(1).

1. Introduction

Europe, it has been suggested, has the strongest non-native speaker (NNS) English proficiency of any region in the world (EF English Proficiency Index, n.d.). English is the language that Europeans are most likely to speak as a second language, with no less than 67% of all respondents to the 2012 Eurobarometer report on Europeans and their Languages indicating that they ‘consider English as one of the two most useful languages for themselves’ (7) and 38% speaking the language well enough to hold a conversation (5). Within Europe, the Netherlands has the greatest proportion of NNSs of English – 90% of the population according to the most recent figures (Eurobarometer, 2012). The country also ranks second in the worldwide English Proficiency Index, with an average *Common European Framework of Reference (CEFR)* level of B2.

The prominent role of English in Dutch society makes Dutch university students of English Language and Culture (ELC) a particularly interesting group of language learners to study. The Oxford Online Placement Test (OOPT) reported on in the previous chapter places the vast majority of ELC students at Radboud University Nijmegen at *CEFR* C1 or even C2 when they come in at the start of year 1. Many have also taken Cambridge Advanced or Proficiency exams in addition to their regular English language classes in secondary school. As these already very advanced language learners are educated in the course of a three-year Bachelor programme to become language professionals, their language needs and goals go over and beyond effective communication. This sets them apart from other language users (De Haan & Van der Haagen, 2012).

Previous studies have suggested that the language production of very advanced language learners is often characterised by a very subtle form of transfer, in which they apply principles of information structure of their L1 to their L2 (cf. Bohnacker & Rosén, 2008; Callies, 2009). In the case of Dutch learners of English

as a Foreign Language (EFL) one of these subtleties seems to be their use of adverbials in pre-subject position, leading to unintentionally marked structures or texts with very heavy-handed discourse linking.

Although the role of transfer in this feature of advanced Dutch learner language is yet to be substantiated within a sound methodological framework (cf. Jarvis, 2000), the studies reported on in the preceding chapters departed from the hypothesis that Dutch learners' frequent use of initial adverbials is likely to result from the difference in information status between clause-initial position in a Subject-Verb-Object (SVO) language like English and a Verb-second (V2) language like Dutch (cf. Verheijen *et al.*, 2013). The V2 principle determines that the finite verb should always be placed in second position in declarative main clauses, whereas the first position is multifunctional and can accommodate subjects, objects and adverbials with equal ease (Table 1):

Table 1. V2 word order

	X	V	S	O	Adv.
a	<i>U</i>	<i>ziet</i>		<i>dat</i>	<i>op uw jaaropgave.</i>
	You	see		that	on your annual statement
b	<i>Dat</i>	<i>ziet</i>	<i>u</i>		<i>op uw jaaropgave.</i>
	That	see	you		on your annual statement
c	<i>Op uw jaaropgave</i>	<i>ziet</i>	<i>u</i>	<i>dat</i>	
	On your annual statement	see	you	that	

SVO word order in English declarative main clauses, on the other hand, is much less flexible. Pre-subject position may but does not need to be occupied and has been argued to be ‘pragmatically marked’ (Los, 2012; Los & Dreschler, 2012; Dreschler, 2015), unmarked links being more likely to be expressed by the grammatical subject of the sentence. This difference in information status seems to come to the surface in Dutch learners’ EFL writing, as in (1) below, where the pre-subject adverbial is probably intended to serve as a neutral discourse link, as it would in Dutch, but is arguably more marked in English:

1. On 29 May 2010, an article was published in the Times called I had my Spider-Man moment. And I failed. **In the article**, Giles Coren narrates about something that happened to him: he witnessed a car accident and, knowing how nasty legal procedures can be, decided to volunteer as a witness.

(student RAD0939, year 1)

The pre-subject place adverbial in bold presents discourse-old information that links back to an antecedent (underlined) in the directly preceding discourse. We will from now on refer to this type of adverbial as a ‘local anchor’ (Los, 2012; Los & Dreschler, 2012).

While place adverbials used for local anchoring stand out, there appear to be differences between Dutch EFL learners and native speakers in the use and relative frequency of other types of adverbials as well. Depending on data-selection, methodology and categorisation of adverbials, advanced Dutch learners appear to use more pronominal adverbs and more pre-subject adverbials overall (Verheijen *et al.*, 2013), more pre-subject place and addition adverbials, including those used for local anchoring (cf. chapter 2), and more local anchors overall, but fewer stance adverbials (cf. chapter 3) than the native-speaker authors of selected

reference corpora. While each of these studies has provided insight into some aspect of Dutch learners' use of pre-subject adverbials, each has a number of inevitable limitations as well. Both Verheijen *et al.* (2013) and the study reported on in chapter 2 demonstrate longitudinal trends in the direction of native writing based on the mean number of adverbials per x number of words at each data collection point. Statistical analysis was not attempted due to the difficulty of dealing with attrition in longitudinal data, with (possibly selective) loss of participants leading to incomplete data sets. The study reported on in chapter 3 did test observed differences between native speakers and Dutch learners for statistical significance. However, as the analysis is cross-sectional rather than longitudinal, it reveals that more proficient learners use fewer local anchors than less proficient learners, but does not trace learners' development over time. All three studies combined also suffer from a lack of comparability due to differences in the selection of learners, reference corpora and text types.

The present study therefore aims to offer a comprehensive overview of PSA use, analysing 780,414 words of written text in total, comprising longitudinal data from four cohorts of Dutch ELC students as well as two NS reference corpora. In order to test whether observed differences between groups and text types represent truly distinct patterns of use, our data will be subjected to statistical analysis.

We will re-examine the question whether or not learners differ from native speakers in their use of PSAs overall, as previous studies have yielded conflicting evidence regarding this point. However, as different classes of adverbials have different preferences for initial, medial and final position, we will also consider whether there are any differences between learners and NSs in the frequency with which they use initial circumstance, stance and linking adverbials. Finally, we will consider the question whether or not learners use more of their

PSAs for local anchoring than native speakers do, as we expect on the basis of previous research. For each of these measures – overall adverbial use, use of separate adverbial classes and use of PSAs for local anchoring – we will not only compare learners and NSs, but also investigate to what extent the frequency with which adverbials are used initially depends on text type, academic discipline and whether or not written assignments were timed or untimed. Finally, we will consider learners' development between year 1 and year 3 of their Bachelor degrees to see whether or not this brings them closer to the NSs.

2. Materials and methods

2.1 Corpus design and annotation

The present study looks at PSA use in a 571,174-word longitudinal corpus of argumentative and literature essays produced by advanced Dutch university students of English, cohorts 2008 up to 2011 of the Dutch component of LONGDALE, the Longitudinal Database of Learner English (Meunier, 2015). We used two NS reference corpora: VUNSPRAC, the VU Native Speaker Published Research Article Corpus (Springer, 2012), consisting of published research articles on literature (70,890 words) and linguistics (79,121 words) written by NS expert writers, and the A-level argumentative essays, written by NS novice writers, included in the Louvain Corpus of Native English Essays (Granger, 1996), amounting to 59,229 words. Table 2 gives an overview of each of the (sub)corpora in terms of the number of texts, contributors and words, as well as the mean word count per text and standard deviation (SD):

Table 2. Corpus composition

			texts	contributors	words	mean	SD
NNS LONGDALE year 1	argumentative	timed	92	92	25,532	277.52	113.59
		untimed	83	83	32,458	391.06	86.86
	literature	untimed	364	203	264,496	726.64	196.34
NNS LONGDALE year 2	argumentative	timed	77	77	22,298	289.58	93.67
		untimed					
	literature	untimed	88	88	178,041	2,023.19	1,444.26
NNS LONGDALE year 3	argumentative	timed					
		untimed					
	literature	untimed	32	32	48,349	1,510.91	606.50
NS LOCNESS	argumentative	timed	110	110	59,229	538.45	181.79
NS VUNSPRAC	linguistics	untimed	11	11	79,121	7,192.82	4,073.95
	literature	untimed	11	11	70,890	6,444.55	3,481.82

The reason for using two NS corpora was that they would provide us with two relevant comparisons: while the novice writers of the NS LOCNESS texts match our students' writing expertise at the start of year one, the NS VUNSPRAC texts represent the kind of expert academic writing that is most relevant to students in the context of their studies. Still, we cannot compare both NS corpora to our entire NNS corpus, as the comparison between EFL learners and novice and expert NS writers would likely be confounded by differences in text type and academic discipline, which we demonstrate in an initial comparison of (1) NS linguistics and literature papers as well as argumentative essays and (2) NNS literature and argumentative essays. In view of the differences between the various text types within the NS and NNS corpora, we did not compare the linguistics component of VUNSPRAC with the NNS LONGDALE corpus, as the latter does not include

linguistics papers. Rather, we compared the NNS argumentative essays with the NS LOCNESS corpus, while the NNS literature essays are only compared with the literature component of the NS VUNSPRAC corpus, bearing in mind that any differences we find in the latter comparison might be partially caused by differences in writing expertise.

Each of the corpora was parsed using the Stanford Parser (Klein & Manning, 2003), after which adverbial phrases occurring in pre-subject position in declarative main clauses were identified with Corpus Studio (Komen, 2012). The output of the Corpus Studio query was a database of 8263 PSAs for NNS LONGDALE, 558 for NS LOCNESS, and 1391 for NS VUNSPRAC. Once the annotation procedure (cf. chapter 1) had been completed, frequencies of each adverbial type were normalised per main clause and subjected to statistical analysis, details of which will be provided below.

2.2 Data analysis

For all of our analyses we used Bayesian statistical methods. As these remain less widely known than the ‘traditional’ frequentist methods developed by Fisher, Pearson and others⁷, we will here present the motivation for our choice and a brief overview of Bayesian parameter estimation, as well as the technical details of the analyses.

2.2.1 Motivation for using Bayesian analysis

We chose a Bayesian analysis for our data in order to be able to measure the conditional probabilities of our hypotheses concerning the relative frequencies of PSAs, given the data obtained from the corpora. This is not possible with

⁷ For an accessible discussion of the differences between frequentist and Bayesian statistics, the reader is referred to Laskin (2016).

frequentist methods, as these do not measure probabilities of hypotheses directly, but rather the conditional probabilities of the data given a null hypothesis. Littré (2014) describes this as one of the most significant differences between frequentist and Bayesian approaches:

[F]requentist results generally tell us how likely it is the data would be observed given some value of θ , while Bayesian results tell us how likely a value of θ is given the observed data. This arguably corresponds better to what researchers are typically most interested in.

Furthermore, due to the large number of comparisons to be made, we wanted to avoid reduction in power of our analysis which is caused by the infamous multiple comparisons problem. Fortunately, Bayesian methods are immune to it (cf. Kruschke, 2010).

2.2.2 Bayesian parameter estimation

Bayesian parameter estimation is based on Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

The $P(A|B)$ notation represents the conditional probability mass or density⁸ of A given B, whereas the $P(A)$ notation stands for the marginal probability mass or

⁸ When estimating the probability distribution of a random variable (i.e. the probabilities associated with possible values the variable can take), what determines if that distribution's associated function is a probability mass or a probability density function is whether the random variable is discrete or continuous. For instance, when rolling a die, or flipping a coin, there is a finite number of discrete possible outcomes, so in that case the y-axis of their distributions will represent probability mass. Probability mass can be thought of as having the same meaning as the term probability. For instance, when flipping a fair coin, the

density of A (i.e. the unconditional probability). In the Bayesian data analysis tradition, the four terms of Eq. 1 are named as follows:

$P(A)$: prior

$P(A|B)$: posterior

$P(B|A)$: likelihood

$P(B)$: evidence.

The prior can be thought of as the probability of, or the degree of belief in, a particular hypothesis, which is not conditional on the obtained data, but only on our previous knowledge (e.g. previous observations or past research related to the phenomenon at hand). The posterior, on the other hand, is the updated prior after taking a particular piece of information into consideration. As can be seen from Eq. 1, this probability is calculated by multiplying the prior by the likelihood term and then dividing by the evidence (called normalisation). An intuitive way of thinking about the likelihood term is as a quantity that represents how much B is consistent with A.

Let us consider an example illustrating the application of Bayes' theorem. A lay meteorologist, armed with some information regarding current weather

probability mass associated with 'heads' and 'tails' is 0.5. Probability densities, on the other hand, are used when dealing with continuous random variables, such as the average height, weight or IQ of a randomly drawn sample of people. In such cases, assigning non-zero probabilities to exact values (with infinite precision) is no longer meaningful and probabilities are instead assigned to intervals of arbitrary length (i.e. to a range of values). Probabilities are calculated by mathematically integrating the probability densities over the respective interval. Probability densities can take any value between 0 and positive infinity, unlike probability masses, whose values are bound to be between 0 and 1. The reason is that, if a lot of probability mass is 'packed' within a small range of values on the x-axis, the values on the y-axis will correspondingly be higher, so that the area under the curve of the distribution remains equal to 1. This allows probability masses to have values between 0 and 1, after being calculated by integrating over any particular range of x-axis values.

conditions, might want to calculate the probability that it will rain on a particular day. Substituting the appropriate terms in equation 1, they would get:

$$P(\text{rain}|\text{weather conditions}) = \frac{P(\text{weather conditions}|\text{rain})P(\text{rain})}{P(\text{weather conditions})} \quad (2)$$

The first step in this calculation would be looking at the data for rain frequency on the same day for the past 100 years and calculating the prior probability $P(\text{rain})$ as the number of rainy days divided by the total number of days (in this case 100). Next, the meteorologist would examine the weather conditions on that particular day (such as the clouds, the wind velocity, and so on) and find whether these conditions are usually associated with a high or low probability of rain (e.g. sunny and quiet weather is much less likely to be followed by rain compared with cloudy and windy weather). This will determine the value for the likelihood term $P(\text{weather conditions}|\text{rain})$. Finally, the evidence $P(\text{weather conditions})$ would be calculated as the total probability of observing the respective conditions (both for days on which rain followed and days on which rain did not follow the same weather conditions).

Equation 1 can be used in estimation theory by treating a parameter as a quantity with a fixed but unknown value and model our uncertainty for it as a random variable. If we denote the parameter with θ , a particular value of the parameter with x , and the collected empirical data with D , Eq. 1 can be rewritten as follows:

$$P(\theta = x|D) = \frac{P(D|\theta = x)P(\theta = x)}{P(D)} \quad (3)$$

In other words, the prior belief in a particular value x of the parameter θ is updated after observing D . Applying Eq. 3 to all values of the parameter gives the full posterior distribution. This is visualised as a curve representing all possible values and their respective probabilities (i.e. the probabilities of these values representing the true value of the parameter).

This approach is fundamentally different from frequentist parameter estimation techniques in that the latter do not directly relate all possible values to their relative probability. Instead, the empirically obtained point estimate is either compared to a particular value and the difference is tested for statistical significance (called null hypothesis significance testing) or, alternatively, a confidence interval is constructed around it.

2.2.3 Drop-outs

Our longitudinal corpus naturally suffers from attrition, as students who contributed towards the corpus at one point, might not necessarily have completed their studies. Those who drop out usually do so in the course of the first year. We felt it was necessary to check if this group was fundamentally different from the remainder of the students who contributed to the corpus, so in order to determine whether or not to exclude those students who dropped out before the end of year 1, we calculated the effect size (also known as Cohen's d) of the factor 'drop-out' for local anchors per main clause. This is given by the following formula:

$$d = \frac{\text{mean}_{\text{nondropouts}} - \text{mean}_{\text{dropouts}}}{\text{std}},$$

$$\text{std} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}},$$

where d is the effect size, std is the pooled standard deviation, m , m , s_1 , and s_2 are the sample sizes and standard deviations of drop-outs and non-drop-outs, respectively. As the resulting effect size was negligibly small ($d = 0.0602$), it was decided not to exclude those students who dropped out within the first year from the main analysis.

2.2.4 Main analysis

We used a Bernoulli likelihood function and a beta conjugate prior with both parameters set equal to 1 (i.e. we used a uniform prior distribution), which allowed us to calculate the posterior distributions for each parameter analytically. The likelihood function was calculated using the empirical frequencies of each adverbial type per total number of main clauses per text for particular groups of participants or text types.

The posterior distributions are extremely rich in information, which can make their interpretation and/or comparison with other posterior distributions difficult. In order to simplify the task, we performed two additional analyses. First, we calculated the mean as a value representing (or summarising) each posterior distribution, by weighing each value on the x-axis by its corresponding probability density. This takes both the most likely value into account, as well as the uncertainty (or variance) of the distribution. Second, when comparing two distributions, we calculated the so-called Probability of Equal Coefficients (PoEC)⁹, which shows the probability that the estimated parameter values for the corresponding two groups are the same.

⁹ A PoEC is obtained by numerically integrating over the product of the two distributions with a fixed step parameter which determines the difference below which two coefficients are accepted as equal. In our case, we used a step value of 0.001.

3. Results

3.1 Text type, discipline and writing expertise in NS writing

A comparison between the three NS corpora confirms that different types of NS writing have distinctly different patterns of PSA use. Not only are there major differences between the LOCNESS argumentative essays and the VUNSPRAC research articles – which might partly be attributed to differences in text type and writing expertise – the two disciplines represented in VUNSPRAC, i.e. literature (lit) and linguistics (ling), also differ greatly between themselves.

3.1.1 Adverbials overall and circumstance, linking and stance adverbials

As Figure 1 shows, the linguistics articles and both other NS corpora have completely non-overlapping distributions of PSA use.¹⁰ Between the literature articles and the LOCNESS essays there is a small degree of overlap.

¹⁰ In each of the plots the parameter being estimated is the probability of a given type of adverbial occurring in a main clause. The x-axis represents a range of possible outcomes, where 0 would mean no use of a given type of adverbial and 1 would mean that each clause starts with that particular type of adverbial. The y-axis gives the corresponding probability density for each possible value given the data.

An easy way of visually obtaining information from a plot of a probability density function is by observing the interval(s) with predominantly high probability densities (relative to other intervals). The actual value of a random variable has a high probability of being within such intervals. Hence, narrow distributions (i.e. ones having low variance) leave less uncertainty regarding the true value of the variable. Furthermore, if two distributions are plotted together, the relative size of the area of overlap between their respective curves is indicative of the probability that the variables associated with the two distributions have the same value. The fact that the distributions of the linguistics articles and other NS corpora in Figure 1 do not overlap is a first indication that the credible values of the linguistics articles are higher than those for the literature articles or LOCNESS essays.

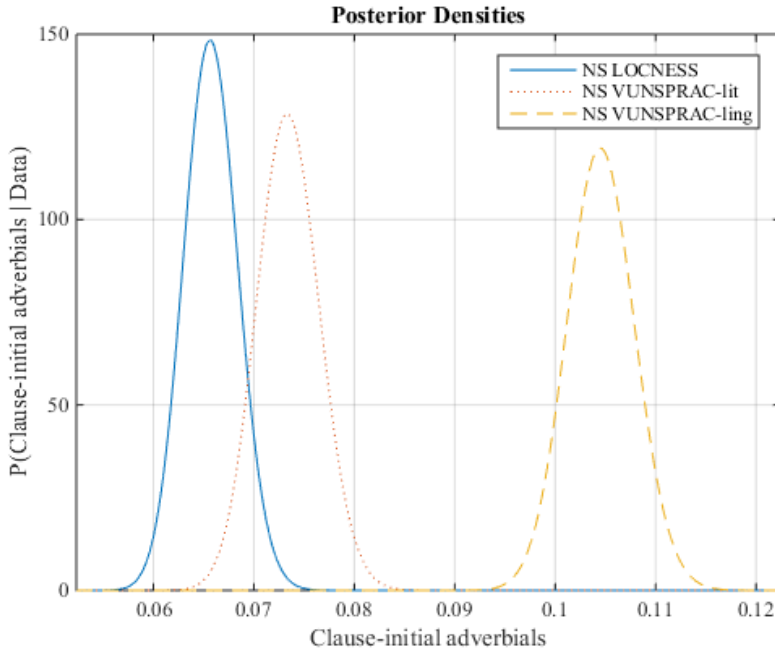


Figure 1. PSAs in LOCNESS, VUNSPRAC-lit and VUNSPRAC-ling

With a mean use per main clause¹¹ of 0.105, the VUNSPRAC-ling articles use considerably more PSAs overall than both the VUNSPRAC-lit articles and the LOCNESS essays, which have means of 0.073 and 0.066 respectively (Figure 2).

¹¹ Throughout the article, 'mean' refers to the mean value of the posterior distribution for a particular group.

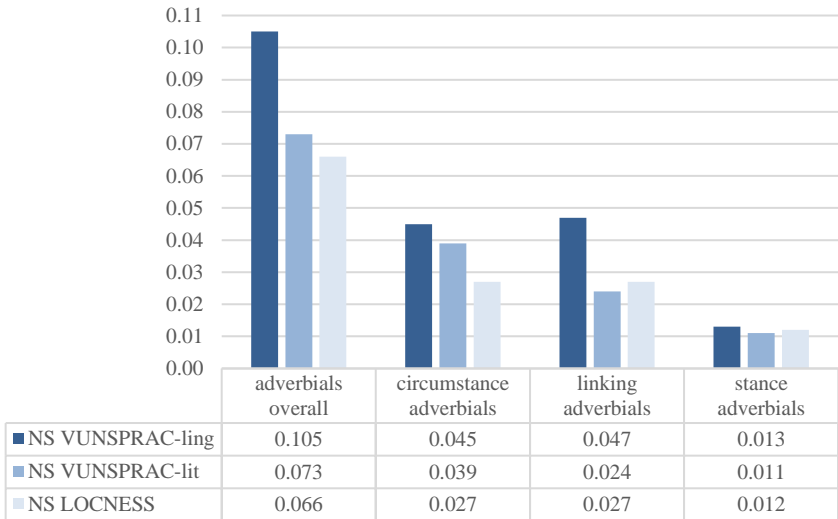


Figure 2. Mean values NS corpora

The differences between the NS corpora are further confirmed by their Probabilities of Equal Coefficients (PoECs). As might be expected on the basis of Figure 1, the linguistics articles have a 0.000 PoEC with both other NS corpora for total use of adverbials in pre-subject position, while the slight overlap between the literature articles and LOCNESS essays is reflected in a somewhat higher PoEC of 0.017¹² (Table 3):

¹² A PoEC of ‘0.000’, as reported here, is understood to mean ‘smaller than 0.001’, as probabilities are only reported up to a certain precision. Which probabilities are small and which probabilities are large is a slightly subjective question. In analogy with traditional p-values, one could decide that if the PoEC is lower than 0.05, it might be reasonable to assume they are not equal. Or alternatively, if the probability is higher than 0.95, it might be decided they are equal. We have decided to keep the original probabilities to allow readers to decide for themselves how strongly to believe in certain hypotheses.

Table 3. PoECs NS corpora

adverbial type	NS VUNSPRAC-ling/ NS VUNSPRAC-lit	NS VUNSPRAC-ling/ NS LOCNESS	NS VUNSPRAC-lit/ NS LOCNESS
adverbials overall	0.000	0.000	0.017
circumstance adverbials	0.021	0.000	0.000
linking adverbials	0.000	0.000	0.095
stance adverbials	0.092	0.205	0.149

The higher mean use of PSAs overall for the linguistics articles is reflected in a higher mean for each of the adverbial classes (Figure 2). There is a slight overlap between the linguistics and literature articles in the distributions for the class of circumstance adverbials, with a 0.021 PoEC between both disciplines (Table 3). The LOCNESS essays include considerably fewer circumstance adverbials, leading to 0.000 PoECs with both other NS corpora. In their use of linking adverbials the LOCNESS essays and literature articles are more alike. At 0.095 their PoEC is relatively high. Distributions for VUNSPRAC-ling and VUNSPRAC-lit, on the other hand, as well as for VUNSPRAC-ling and LOCNESS are completely non-overlapping, with a 0.000 PoEC. Differences in the use of stance adverbials, finally, are small, leading to a relatively high PoEC of 0.092 between VUNSPRAC-ling and VUNSPRAC-lit, 0.205 between VUNSPRAC-ling and LOCNESS and 0.149 between VUNSPRAC-lit and LOCNESS.

3.1.2 *Local anchors*

The linguistics articles also feature more local anchors than the other NS corpora, with a mean use of 0.006 per main clause for VUNSPRAC-ling and 0.005 and 0.002 for VUNSPRAC-lit and LOCNESS respectively (Figure 3).

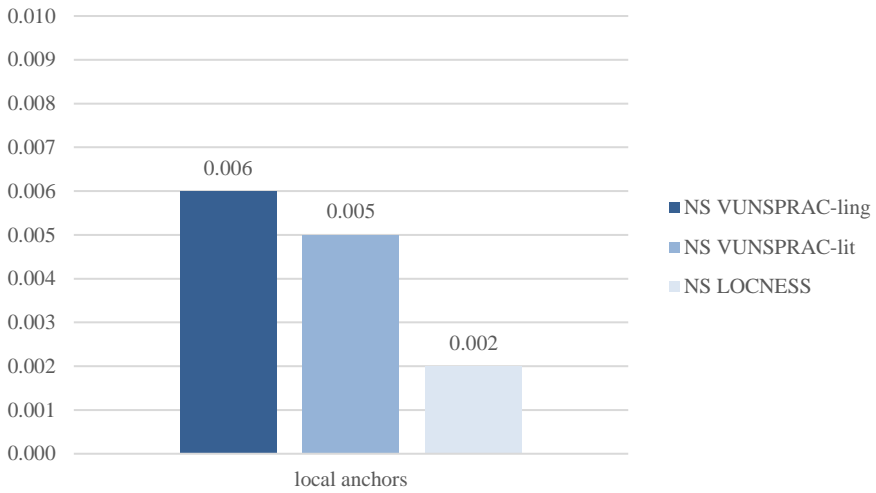


Figure 3. Mean values NS corpora

The difference between VUNSPRAC-ling and VUNSPRAC-lit is small, with a 0.312 PoEC (Table 4), and considerable overlap between the distributions of both disciplines (Figure 4).

Table 4. PoECs NS corpora

adverbial type	NS VUNSPRAC-ling/ NS VUNSPRAC-lit	NS VUNSPRAC-ling/ NS LOCNESS	NS VUNSPRAC-lit/ NS LOCNESS
local anchors	0.312	0.002	0.004

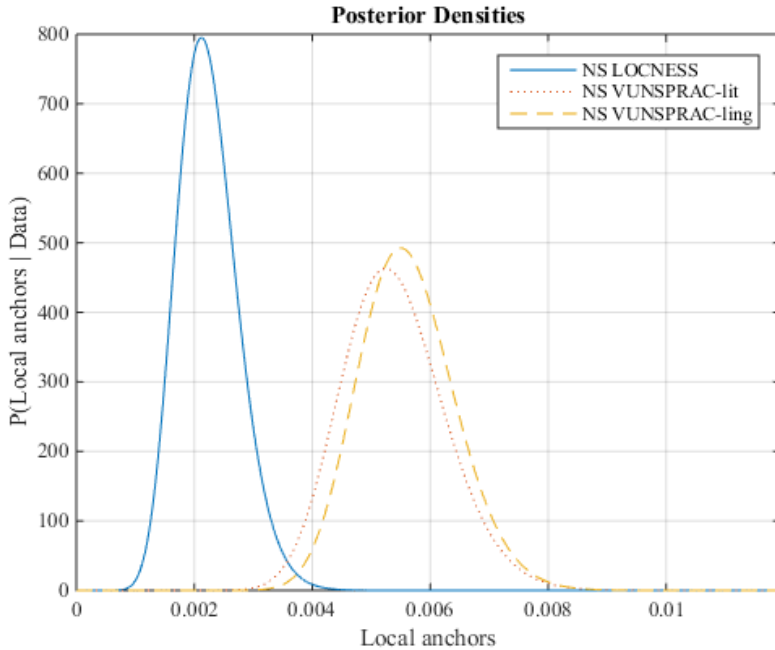


Figure 4. Local anchors in NS LOCNESS, VUNSPRAC-lit and VUNSPRAC-ling

LOCNESS, on the other hand, has an almost completely non-overlapping distribution with both the linguistics and the literature articles, with PoECs of 0.002 and 0.004 with VUNSPRAC-ling and VUNSPRAC-lit respectively. It is very rare for the authors of the LOCNESS argumentative essays to use adverbials anchored in the preceding sentence as a point of departure for the subsequent discourse. In the VUNSPRAC corpus this is a relatively more common cohesive strategy, especially in contexts where there is either contrast (example 2) or a shift in perspective (example 3):

2. Like the earlier adventure tales, cybernarratives project an imperialist organization of the world in their mapping of the physical world, as their real spaces rely on the division between first and third worlds, demarcated by race and ethnicity. **In cyberfiction**, however, the imperialism is Americanized cultural domination rather than European governmental colonization.

(NS lit_07)

3. While the AP and Hansard corpora display a much higher percentage of modifier cases than of head cases, this trend is not so marked in the APHB Corpus, where there are only slightly more modifier cases, especially with this and that, as we see in Table 9. **Within the modifier cases**, it is worth noting a distinction that can be made between modifier-demonstratives with head words that provide new information about their antecedents (examples 9 and 10 below) and those that simply corefer with their antecedents, without any new information being provided.

(NS ling_04)

It appears that the use of initial circumstance adverbials for local anchoring is a feature that occurs relatively more frequently in formal academic writing, both in the literature and in the linguistics articles.

3.2 Argumentative and literature essays in the Dutch component of LONGDALE

3.2.1 Adverbials overall and circumstance, linking and stance adverbials

The NNS argumentative essays (LONGDALE-arg) include more PSAs than the literature essays (LONGDALE-lit). This is mainly due to a considerably higher mean use of linking adverbials of 0.040 in LONGDALE-arg, against 0.025 in LONGDALE-lit (Figure 5).

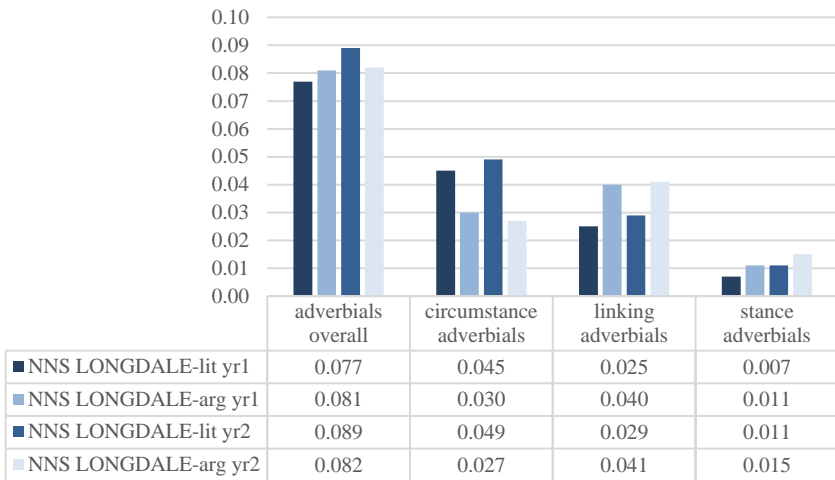


Figure 5. Mean values NNS LONGDALE-lit and LONGDALE-arg

Circumstance adverbials feature more heavily in the literature essays, with a mean use of 0.045 against 0.030 for the argumentative essays, which may be explained by the frequent use of place adverbials to refer to the text that is being discussed, such as ‘in this poem’, ‘in chapter three’ and ‘in his early work’ (cf. chapter 2). At 0.011 against 0.007, mean use of stance adverbials, finally, is slightly higher in the argumentative essays than in the literature essays. These distinct patterns of distribution mirror the differences between the NS corpora, where the LOCNESS argumentative essays also included more linking adverbials and slightly more

stance adverbials, while the VUNSPRAC-lit articles used more circumstance adverbials.

PoECs between both text types in the NNS corpus are generally low. In year 1 the total number of adverbials does not differ greatly, as a higher use of one adverbial class compensates for the lower use of another, so that the PoEC of 0.060 is relatively high, especially compared with the 0.003 PoEC for stance adverbials and 0.000 for both circumstance and linking adverbials (Table 5).

Table 5. PoECs NNS LONGDALE-lit/LONGDALE-arg

adverbial type	NNS LONGDALE-lit/ NNS LONGDALE-arg year 1	NNS LONGDALE-lit/ NNS LONGDALE-arg year 2
adverbials overall	0.060	0.028
circumstance adverbials	0.000	0.000
linking adverbials	0.000	0.000
stance adverbials	0.003	0.050

In year 2, distributions for circumstance and linking adverbials are still completely non-overlapping, with 0.000 PoECs between both text types, while the PoEC for the class of stance adverbials has increased slightly to 0.050, still low enough to assume that argumentative and literature essays have distinct patterns of distribution. The PoEC for the total number of PSAs, meanwhile, has decreased to 0.028.

3.2.2. Local anchors

The NNS literature essays include more local anchors than the argumentative essays, both in year 1 and in year 2 (Figure 6), mirroring the difference in local anchoring between the NS literature articles and argumentative essays.

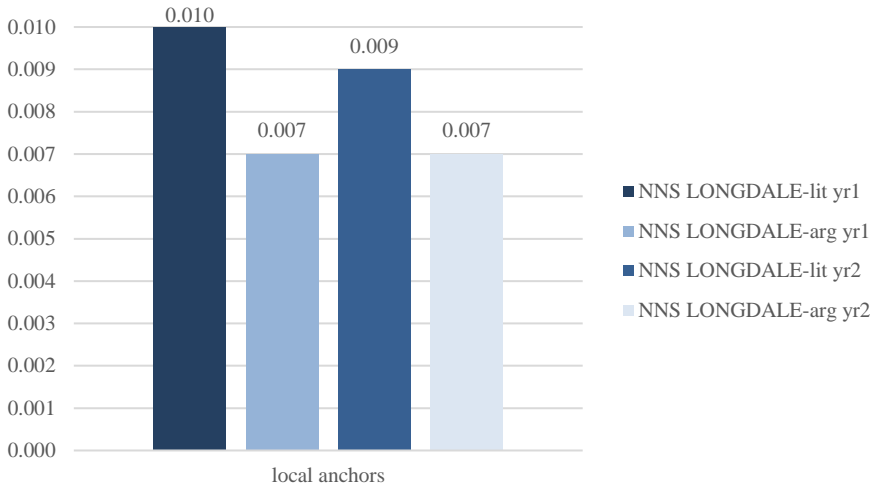


Figure 6. Mean values LONGDALE-lit and LONGDALE-arg

Due to a decrease in the number of local anchors in the literature essays, the difference between literature and argumentative essays is somewhat smaller in year 2, leading to a rise in the PoEC from 0.011 in year 1 to 0.084 in year 2 (Table 6).

Table 6. PoECs NNS LONGDALE-lit/ LONGDALE-arg

adverbial type	NNS LONGDALE-lit/ NNS LONGDALE-arg year 1	NNS LONGDALE-lit/ NNS LONGDALE-arg year 2
local anchors	0.011	0.084

3.3 Novice NNS and expert NS literature papers: LONGDALE-lit vs VUNSPRAC-lit

3.3.1 *Adverbials overall and circumstance, linking and stance adverbials*

In year 1 the NNSs use slightly more adverbials overall in their literature essays than the authors of the most comparable NS corpus, the VUNSPRAC-lit articles. This can be attributed to a higher use of circumstance and linking adverbials:

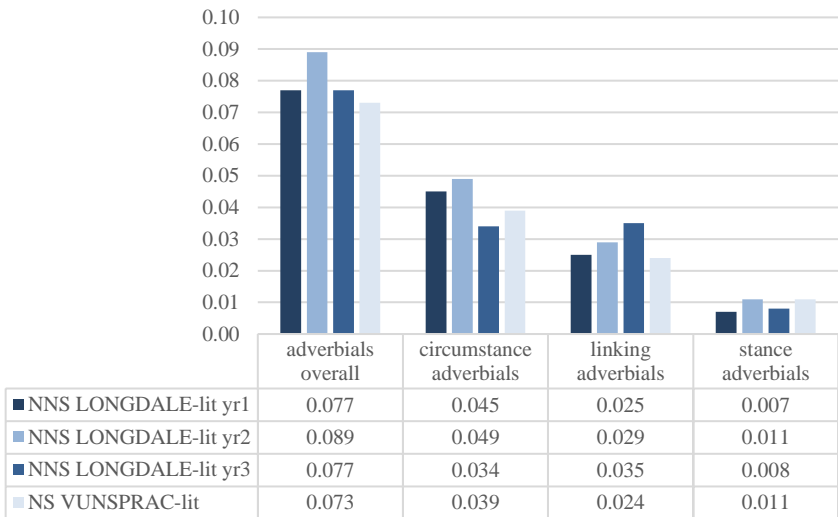


Figure 7. Mean values NNS LONGDALE-lit/ NS VUNSPRAC-lit

The only adverbial class which the NSs use more of is that of stance adverbials. The kind of explicit commentary on the proposition in the remainder of the clause that is conveyed by the stance adverbials exemplified in the sentences below is quite rare in the NNS literature essays:

4. Yet, **conceivably**, what the novel of ideas represents is the commodification or the reification – inadvertent, inevitable, or otherwise – of the philosophical novel.

(NS lit_01)

5. **Significantly**, McGann begins his book – which features treatments of Byron, Dante Gabriel Rossetti, and Ezra Pound – with an attempt to show through Blake “how poetry is a form of action rather than a form of representation”.

(NS lit_03)

6. **Ironically**, in flaunting new possibilities of distance-transcending technology, these narratives rehearse old geographic interpretations of space.

(NS lit_07)

In spite of the higher mean value, there is considerable overlap between NNSs and NSs in the use of adverbials overall and, more particularly, in the use of linking adverbials, with PoECs of 0.069 and 0.189 respectively (Table 7). The lower PoECs of 0.008 for stance adverbials and 0.009 for circumstance adverbials suggest that there might be a reliable difference between the two groups in the use of these two adverbial classes in year 1.

Table 7. PoECs NNS LONGDALE-lit/ NS VUNSPRAC-lit

adverbial type	NNS LONGDALE-lit year 1/ NS VUNSPRAC-lit	NNS LONGDALE-lit year 2/ NS VUNSPRAC-lit	NNS LONGDALE-lit year 3/ NS VUNSPRAC-lit
adverbials overall	0.069	0.000	0.062
circumstance adverbials	0.009	0.000	0.047
linking adverbials	0.189	0.009	0.000
stance adverbials	0.008	0.259	0.050

Perhaps surprisingly, the mean frequency of each adverbial class increases between year 1 and 2, widening the gap between NSs and NNSs, except in the case of stance adverbials, where the increase represents a trend in the direction of NS writing, leading to a 0.259 PoEC. For each of the other adverbial classes, PoECs between NSs and NNSs go down, to 0.009 in the case of linking adverbials, and to 0.000 for both circumstance adverbials and adverbials overall. PoECs between year 1 and 2 within LONGDALE-lit are invariably low (Table 8, column 1).

Table 8. PoECs NNS LONGDALE-lit

adverbial type	NNS LONGDALE-lit year 1/year 2	NNS LONGDALE-lit year 2/year 3	NNS LONGDALE-lit year 1/year 3
adverbials overall	0.000	0.001	0.109
circumstance adverbials	0.028	0.000	0.000
linking adverbials	0.001	0.011	0.000
stance adverbials	0.000	0.016	0.302

Between year 2 and 3 the use of adverbials overall decreases again, returning to the year 1 starting level. The year 1 and year 3 distributions most closely resemble the NS VUNSPRAC-lit distribution, with a 0.069 and 0.062 PoEC respectively, compared with 0.000 for year 2 (Table 7). As the use of adverbials overall ends up close to its starting point, the PoEC between year 1 and 3 is relatively high, at 0.109 (compared with 0.000 for year 1 and 2 and 0.001 for year 2 and 3, see Table 8). This does not mean to say, however, that there is no development in the NNS literature essays between year 1 and 3, as different adverbial classes show widely differing patterns so that the rise in the use of one class is compensated by the decline in another.

Breaking down the use of adverbials overall into separate adverbial classes, we observe a clear development in the use of circumstance adverbials in the direction of the NSs. After an initial rise between year 1 and 2, the use of circumstance adverbials then drops below the year 1 starting level (Figure 8), leading to a rise in the PoEC between NSs and NNSs to 0.047 by year 3 (Table 7).

There is less overlap within the NNS LONGDALE-lit corpus: the year 3 distribution has a 0.000 PoEC with both year 1 and year 2 (Table 8).

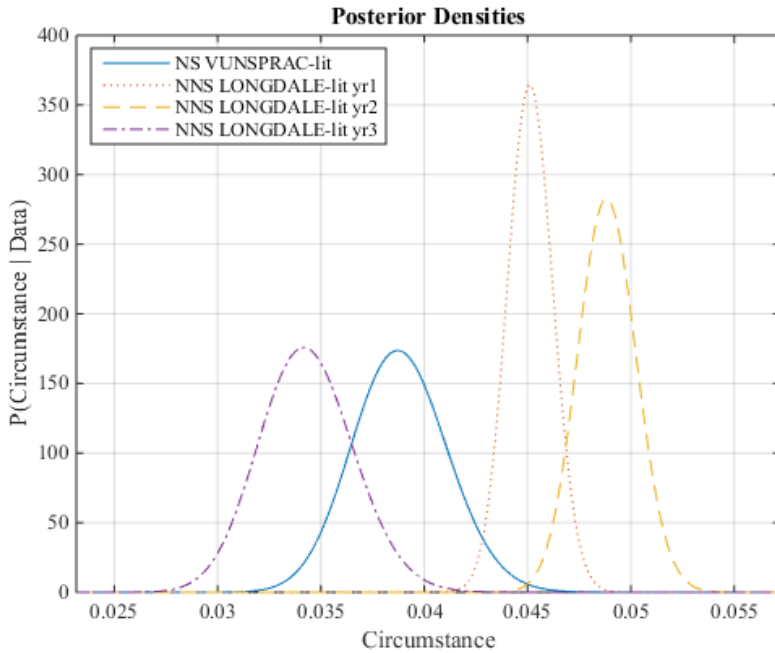


Figure 8. Circumstance adverbials per main clause in NNS LONGDALE-lit year 1, 2 and 3 and NS VUNSPRAC-lit

Meanwhile, the number of linking adverbials continues to rise, moving further and further away from the mean of the NSs (Figure 9). While the NNS and NS distributions showed considerable overlap in year 1, the continued increase in their use by the NNSs leads to completely non-overlapping distributions by year 3, with a 0.000 PoEC (Table 7). Probabilities between successive years within the NNS corpus are low as well (Table 8), demonstrating a clear longitudinal development that is indicative of an unexpectedly widening gap between NSs and NNSs.

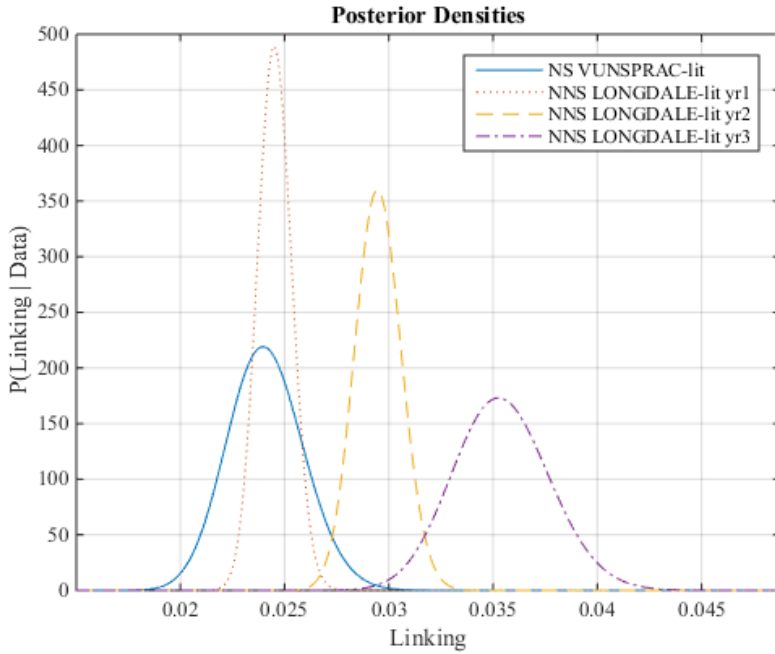


Figure 9. Linking adverbials in NNS LONGDALE-lit and NS VUNSPRAC-lit

The class of stance adverbials is the only one that merely fluctuates between year 1 and 3. Use of stance adverbials had approached the mean of the NS VUNSPRAC-lit texts in year 2, but in year 3 it drops below the NS level again, so that the PoEC between NSs and NNSs decreases to 0.050. Probabilities between successive years in the NNS LONGDALE-lit corpus follow the same pattern as use of adverbials overall: low PoECs of 0.000 and 0.016 between year 1 and 2, and year 2 and 3 respectively, but considerable overlap between year 1 and 3 with a 0.302 PoEC (Table 8).

3.3.2. *Local anchors*

As expected, use of local anchors in NNS LONGDALE-lit exceeds use in NS VUNSPRAC-lit. Local anchoring in LONGDALE-lit decreases slowly but steadily between year 1 and 3, from 0.010 in year 1, to 0.009 in year 2 and 0.008 in year 3:

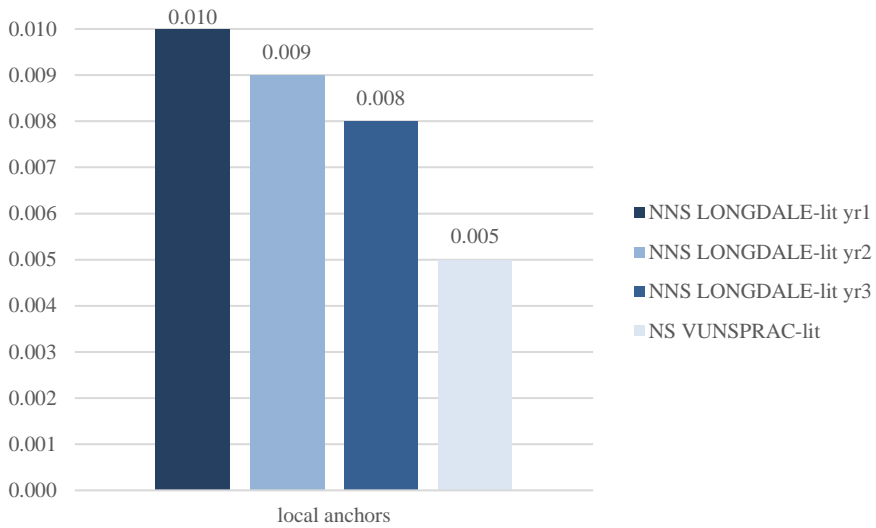


Figure 10. Mean values NNS LONGDALE-lit/ NS VUNSPRAC-lit

While a difference of 0.001 between successive years is more meaningful for the relatively small category of local anchors than it would be for adverbials overall, there is still considerable overlap between the posterior distributions, and PoECs are relatively high (Table 9).

Table 9. PoECs NNS LONGDALE-lit

adverbial type	NNS LONGDALE-lit year 1/year 2	NNS LONGDALE-lit year 2/year 3	NNS LONGDALE-lit year 1/year 3
local anchors	0.301	0.185	0.085

Table 10. PoECs NNS LONGDALE-lit/NS VUNSPRAC-lit

adverbial type	NNS LONGDALE-lit year 1/ NS VUNSPRAC-lit	NNS LONGDALE-lit year 2/ NS VUNSPRAC-lit	NNS LONGDALE-lit year 3/ NS VUNSPRAC-lit
local anchors	0.000	0.005	0.082

Initially, the NS and NNS distributions are completely non-overlapping, with a 0.000 PoEC (Table 10). As time goes by, however, the NNSs edge closer and closer to the NS mean of 0.005, with higher probabilities and increasing overlap between NS and NNS distributions as a result, as visualised in Figure 11:

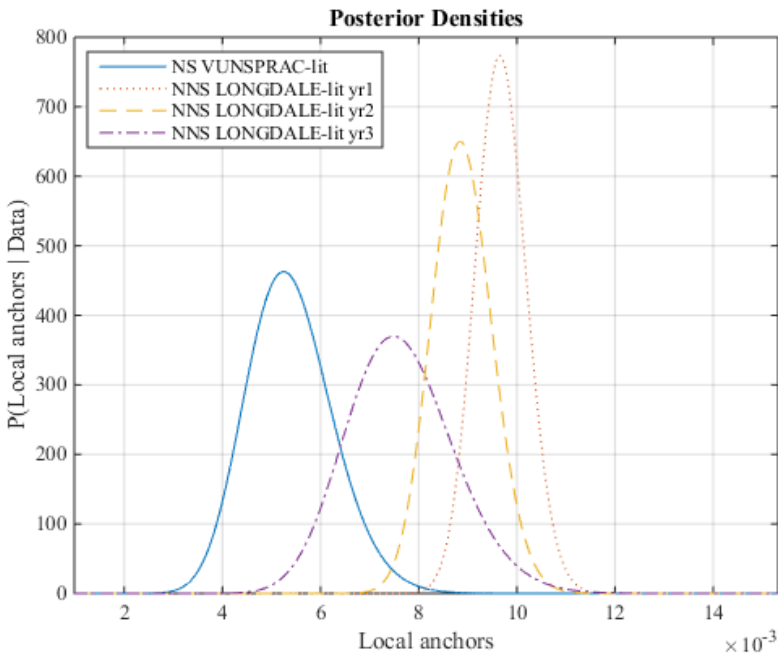


Figure 11. Local anchors in NNS LONGDALE-lit and NS VUNSPRAC-lit

For NSs and NNSs alike, local anchors are overwhelmingly realised as circumstance adverbials, such as the following:

7. The point of Georg's account of the friend is to show, not just the existence of the friend, but the importance of affinity itself. **In his account**, Georg is especially proud that the father began to "get on with" the friend, even to the point of repeating the friend 's story.

(NS lit_04)

8. The fact that the last two lines differ in rhyme scheme from the previous lines makes it even clearer that it is the end of the poem, an, in this poem, also the end of the man. **In the poem**, there is some alliteration (...).

(student RAD1102, year 1)

While it is not unusual for the NNSs to also use linking adverbials for local anchoring, a similar instance of an addition adverbial with pronominal reference to the preceding context occurs only once in all of the NS literature articles:

9. This will be discussed at length in the next section. **Aside from that**, they are also a binary categorization of everything, a pair of polar opposites dividing existence in to the mythic or the real.

(NS lit_06)

3.4 Argumentative writing by novice NS and NNS writers: LONGDALE-arg vs LOCNESS

3.4.1 Adverbials overall and circumstance, linking and stance adverbials

The Dutch authors of the LONGDALE-arg sub-corpus use considerably more PSAs in their argumentative essays than the NS authors of the LOCNESS argumentative essays. Most of these are linking adverbials, which initially have a mean of 0.044, followed by circumstance adverbials with 0.034. In LOCNESS these two adverbial classes are represented equally, with means of 0.027 for both (Figure 12). Stance adverbials are used much less frequently in both corpora: 0.009 per main clause in LONGDALE-arg compared with 0.012 in LOCNESS.

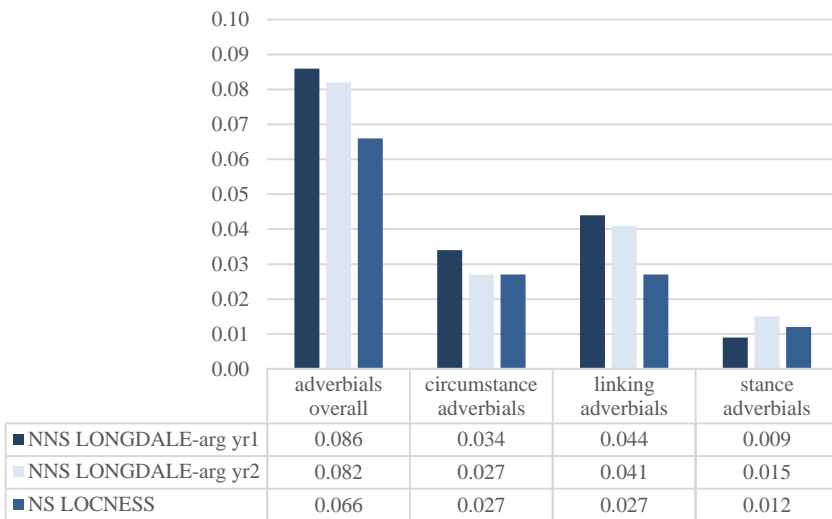


Figure 12. Mean values NNS LONGDALE-arg/ NS LOCNESS argumentative essays

Between year 1 and 2 there is a clear development in the direction of the NNSs. The use of all types of adverbials in initial position decreases, with the exception of stance adverbials, which are again marked by an increase: the NNSs initially use fewer stance adverbials than the NSs but then rise to a mean of 0.015. The decrease

in the use of circumstance adverbials between year 1 and 2 brings the NNSs in line with the NS mean of 0.027. Interestingly, the use of linking adverbials also decreases slightly, from a mean of 0.044 to 0.041, in spite of the fact that in the NNS literature essays the use of the same adverbial class was marked by a steady increase, from a much lower starting level of 0.025 in year 1 to 0.035 in year 3. However, even at its maximum, in year 3, the mean of the literature essays is still some way below the mean of the argumentative essays at any point in time.

The PoECs in Table 11 show that, in spite of some overlap, there is a clear development between year 1 and 2, particularly for the classes of circumstance and stance adverbials.

Table 11. PoECs NNS LONGDALE-arg

adverbial type	NNS LONGDALE-arg year 1/year 2
adverbials overall	0.051
circumstance adverbials	0.030
linking adverbials	0.068
stance adverbials	0.016

Table 12. PoECs NNS LONGDALE-arg/NS LOCNESS

adverbial type	NNS LONGDALE-arg year 1/ NS LOCNESS	NNS LONGDALE-arg year 2/ NS LOCNESS
adverbials overall	0.000	0.001
circumstance adverbials	0.018	0.121
linking adverbials	0.000	0.000
stance adverbials	0.056	0.108

By year 2 the NNSs have moved some way in the direction of the NSs, as is evident from the PoECs listed in Table 12. However, a gap remains as there is still only very slight overlap in the use of adverbials overall, with a PoEC of 0.001, compared with 0.000 for year 1, and the distributions for the class of linking adverbials are still completely non-overlapping by year 2. PoECs for the classes of circumstance and stance adverbials do increase considerably between year 1 and 2, from 0.018 to 0.121 for the circumstance adverbials and from 0.056 to 0.108 for the stance adverbials.

3.4.2 Local anchors

Again, use of local anchors in LONGDALE-arg exceeds the use by NSs in the LOCNESS argumentative essays, with means of 0.009 for the NNSs in year 1 against 0.002 for the NSs (Figure 13). By year 2 NNSs have moved some way in the direction of the NSs, as the use of local anchors decreases to 0.007. A considerable gap remains, however.

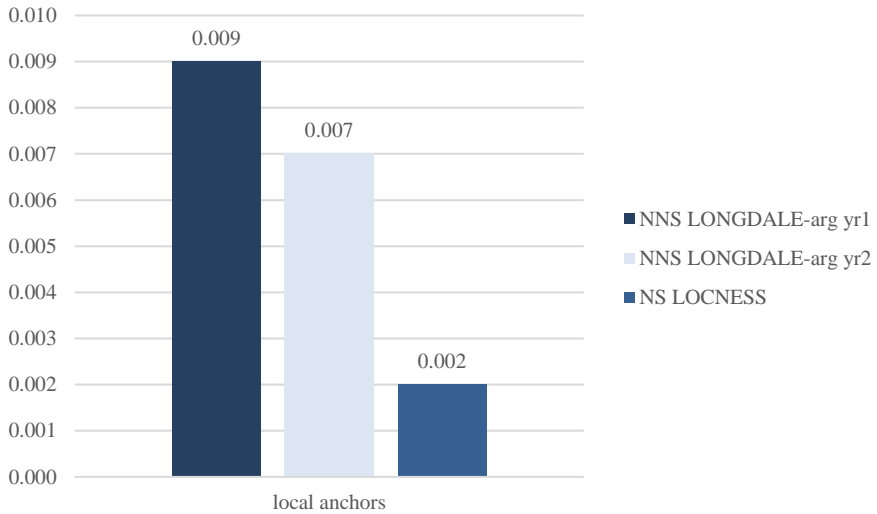


Figure 13. Mean values NNS LONGDALE-arg/NS LOCNESS

The gap between NSs and NNSs, both in year 1 and in year 2, is wider than between both years within LONGDALE-arg. In year 1 NSs and NNSs have a 0.000 PoEC for the use of local anchors overall (Table 14). By year 2 this has increased only slightly, to 0.002.

Table 13. PoEC NNS LONGDALE-arg

adverbial type	NNS LONGDALE-arg year 1/year 2
local anchors	0.088

Table 14. PoEC NNS LONGDALE-arg/NS LOCNESS

adverbial type	NNS LONGDALE-arg year 1/ NS LOCNESS	NNS LONGDALE-arg year 2/ NS LOCNESS
local anchors	0.000	0.002

Apart from differences in relative frequencies, NSs and NNSs also differ in the way local anchors are realised. All of the local anchors in LOCNESS fall into the class of circumstance adverbials. While circumstance adverbials account for the majority of the local anchors used in LONGDALE-arg as well, the argumentative essays produced by the NNSs also feature a small number of stance (10) and a slightly larger number of linking adverbials (11) that link back to an antecedent in the preceding discourse:

10. As was said in the given topic statement, a gap year provides a challenge that builds character and increases independence, and for some this will certainly be true. **For them**, it can be helpful and beneficial, but what is good for a few should not be used as advice for all.

(student RAD1039, year 2)

11. By describing critical thoughts such as these the audience will notice that your writing is objective. **On top of that**, your reliability might improve when you doubt your own opinions or conclusions.

(student RAD1059, year 1)

3.5 Timed vs untimed NNS writing

3.5.1 Adverbials overall and circumstance, linking and stance adverbials

Apart from looking at development longitudinally, the inclusion in the LONGDALE-arg corpus of timed as well as untimed year 1 argumentative essays also presents an opportunity to investigate whether the circumstances under which an assignment is written have an effect on adverbial use. Interestingly, for each of the adverbial classes, untimed essays are closer to the NS mean than timed essays. Figure 14 shows that untimed essays use considerably fewer adverbials overall, fewer circumstance adverbials and fewer linking adverbials. Once again, the opposite is true for the class of stance adverbials, with a higher mean in the untimed compared with the timed essays. The untimed mean for stance adverbials is equal to that of the NS LOCNESS essays.

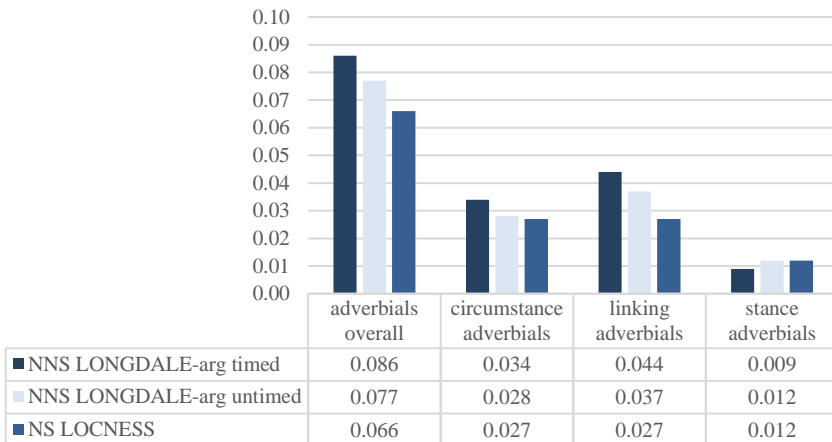


Figure 14. Mean values NNS LONGDALE-arg timed/untimed and NS LOCNESS

The more native-like use of PSAs in untimed compared with timed essays is confirmed by the PoECs listed in Table 15 below, with the exception of linking adverbials. Although the untimed essays are slightly closer to the NS LOCNESS

distribution than the timed essays (Figure 15), the frequency of initial linking adverbials is considerably higher in both timed and untimed essays than it is in LOCNESS, leading to a 0.000 PoEC in both cases.

Table 15. PoECs NNS LONGDALE-arg timed/ untimed and NS LOCNESS

adverbial type	NNS LONGDALE-arg timed/ NS LOCNESS	NNS LONGDALE-arg untimed/ NS LOCNESS
adverbials overall	0.000	0.003
circumstance adverbials	0.018	0.128
linking adverbials	0.000	0.000
stance adverbials	0.056	0.205

In all other cases, untimed essays have a higher PoEC with LOCNESS than timed essays. For the use of PSAs overall, the PoEC with LOCNESS is only slightly higher for the untimed compared with the timed essays, at 0.003 against 0.000. For the remaining adverbial classes it would be difficult to maintain that there is a reliable difference between LOCNESS and the untimed argumentative essays, while the distance between LOCNESS and the timed essays is much more considerable.

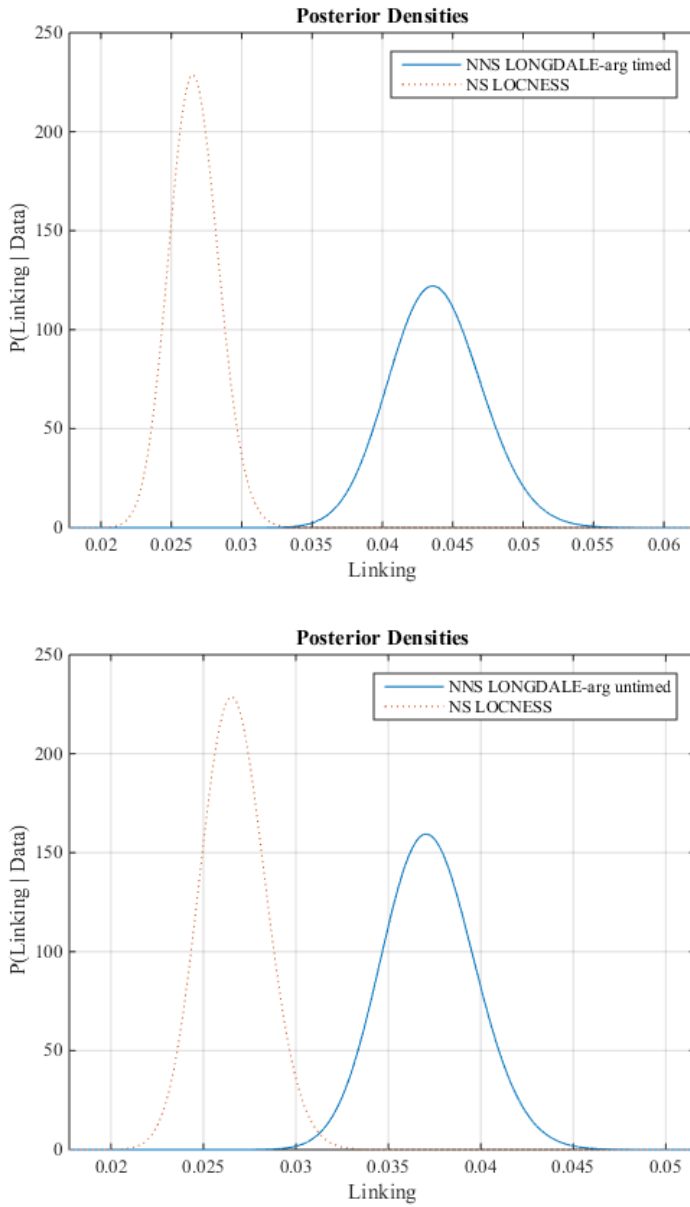


Figure 15. Linking adverbials in NNS LONGDALE-arg timed/untimed and NS LOCNESS

Table 16 also confirms that timed and untimed essays have distinct distributions, with low PoECs between both conditions, with the exception of the class of stance adverbials, for which differences are smaller and PoECs relatively higher.

Table 16. PoECs NNS LONGDALE-arg timed/untimed

adverbial type	NNS LONGDALE-arg timed/untimed
adverbials overall	0.021
circumstance adverbials	0.038
linking adverbials	0.027
stance adverbials	0.074

3.5.2 Local anchors

Untimed essays are closer to LOCNESS than timed essays in their use of local anchors as well, as can be seen in Figure 16:

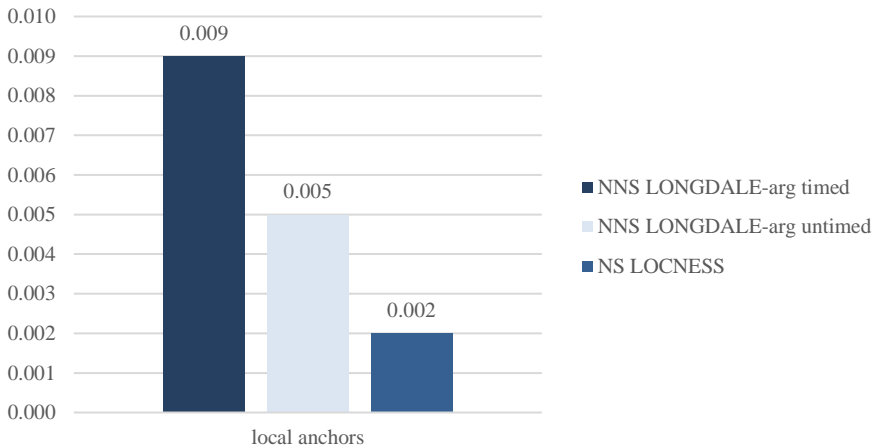


Figure 16. Mean values NNS LONGDALE-arg timed/untimed and NS LOCNESS

PoECs between timed and untimed essays suggest that they have distinct distributions for the use of local anchors (Table 17). PoECs between LOCNESS and untimed essays turn out to be higher than those for LOCNESS and timed essays (Table 18).

Table 17. PoEC NNS LONGDALE-arg timed/untimed

adverbial type	NS LONGDALE-arg timed/untimed
local anchors	0.020

Table 18. PoEC NNS LONGDALE-arg timed/untimed and NS LOCNESS

adverbial type	NNS LONGDALE-arg timed/ NS LOCNESS	NNS LONGDALE-arg untimed/ NS LOCNESS
local anchors	0.000	0.008

Untimed essays are not only more native-like than timed essays in the frequency with which they use local anchors, but also in the way local anchors are realised, featuring fewer local anchors that fall into the class of linking adverbials.

4. Discussion

We have presented a comprehensive overview of use of PSAs in argumentative and literary essays produced by advanced Dutch EFL learners at different stages in their development, as well as in the NS argumentative essays and linguistics and literature articles that were used as a reference. Our results show that the frequency with which adverbials are used clause-initially does indeed depend on the author's L1.

For each of the text types considered, the Dutch learners tend to use more PSAs than the NSs, but as different classes of adverbials have different preferences for initial, medial or final position it is more informative to consider each in

isolation. Learners start out using more circumstance adverbials and linking adverbials, but fewer stance adverbials than NSs. Without exception, the use of local anchors in the LONGDALE corpus also exceeds the use in the NS corpora.

For the learners, PSA use is clearly also affected by longitudinal development. In spite of a peak in year 2, the use of circumstance adverbials in the literature essays develops in the direction of native writing, eventually dropping below the number used in the VUNSPRAC literature articles. Similarly, the number of circumstance adverbials used in the argumentative essays decreases between year 1 and 2 to the exact same frequency found in the LOCNESS argumentative essays. For both text types, the PoEC between NS and NNS corpus is considerably higher in the final year than in the preceding year(s).

The use of linking adverbials in the NNS literature essays develops in the opposite direction. While the year 1 mean is still quite close to the mean of the NS literature articles, the number of linking adverbials continues to rise until the PoEC between the NSs and NNSs is eventually reduced to 0.000. The number of linking adverbials in the NNS argumentative essays does drop slightly between year 1 and 2. However, compared with the literature essays, the year 1 mean of the argumentative essays was much higher to start with and even in year 2 the distance between the NSs and NNSs remains considerable: the PoEC between LONGDALE-arg and LOCNESS remains stable at 0.000.

The class of stance adverbials is the smallest of the three, across corpora and text types. Learners tend to use fewer than native speakers but frequencies fluctuate. Learners do use more stance adverbials in year 2 than they do in year 1, both in the literature and argumentative essays, thereby increasing PoECs between NSs and NNSs, although this effect is countered by the subsequent decrease in the literature essays between year 2 and 3.

PSAs that are used for local anchoring, finally, show a steady decrease between subsequent years in both text types included in the learner corpus. Initially, the distributions of the NS and NNS corpora are completely non-overlapping, with 0.000 PoECs. Probabilities do increase slightly in subsequent years, but a considerable gap remains.

A final factor in the use of PSAs by the Dutch learners is the added pressure of a time-limit. The comparison of timed versus untimed argumentative year 1 essays reveals that untimed essays are more native-like in their use of each of the different types of adverbials considered in this study. In fact, in many cases the distinction between timed and untimed essays is even bigger than between timed essays in year 1 versus year 2, suggesting that it is not just increased knowledge of or exposure to English which makes learners more native-like. It seems plausible that learners already have some awareness of the differences between Dutch and English in the way PSAs are used for discourse linking even in year 1, but under time pressure they will fall back on the default strategies that they are familiar with from their L1.

Other factors influencing the use of PSAs manifest themselves in a comparison between the NS corpora. There is a clear effect of academic discipline, even in the absence of differences in L1 or writing expertise, in that the VUNSPRAC linguistics articles include significantly more linking and circumstance adverbials than the VUNSPRAC literature articles. And while the differences between VUNSPRAC-lit and LOCNESS could be down either to differences in text type or writing expertise, the fact that they are mirrored by the differences between argumentative and literature essays in the NNS corpus seems to confirm that the discourse linking strategies that writers employ depend to a great extent on text type. In the NS and NNS corpora alike, linking adverbials

predominate in the argumentative essays, while circumstance adverbials are used more often in the literature essays and articles.

These results confirm and extend the preliminary conclusions of chapter 2, which looked at the use of place and addition adverbials, the two most frequently used adverbial categories, reporting a decline in the use of place adverbials but an increase in the use of addition adverbials. The present study found that this pattern of use can be extended to the adverbial classes that place and addition adverbials belong to, the classes of circumstance adverbials and linking adverbials respectively. Similarly, the decline in the use of place and addition adverbials used for local anchoring has here been shown to apply to all PSAs used for local anchoring.

Chapter 3 also found a statistically significant higher use of local anchors for Dutch EFL learners compared with native speakers of English. Unlike the present study, however, it did not find a significant difference between Dutch EFL learners and native speakers of English in the use of PSAs overall, nor in their use of linking adverbials. A possible explanation lies in the fact that the study reported on in chapter 3 used a different, *CEFR*-aligned dataset, consisting of untimed texts, particularly because the present study found a relatively smaller difference between learners and native speakers for untimed compared with timed argumentative essays for each of the adverbial classes. The relatively high PoECs between NNS untimed texts and NS LOGNESS essays for the classes of circumstance and linking adverbials do seem to confirm the absence of any statistically significant difference between the groups for these particular adverbial classes in chapter 3, once again highlighting the importance of taking both text type and task instructions into account in corpus construction.

5. Conclusion

Any longitudinal corpus of language produced by university students of a second language is likely to bear witness to a combination of academic and linguistic development. This makes it difficult to tease apart the relative importance of the authors' L1 versus differences in writing expertise in the comparison between the literature essays written by the learners and the published literature articles used as a reference. A comparison with NSs with similar age and academic background included in this study, i.e. the LOCNESS argumentative essays, would be confounded by differences in text type. Use of additional reference corpora could help to address this problem. At this stage, we can at least observe that in year 3 learners' literature essays are generally closer to the NS literature articles in their use of PSAs than they were in year 1, which is likely to be the combined effect of an increase in proficiency and writing expertise.

While this is true for most adverbial classes, as we have seen, the use of linking adverbials is the exception, as the learners' increasing use of adverbials belonging to this class in the literature essays takes them further and further away from the NS mean. Similarly, it is also the class of linking adverbials for which there is least overlap in the comparison of the LOCNESS argumentative essays and the argumentative essays produced by the Dutch learners, both in year 1 and in year 2. Here, the higher mean use by the Dutch learners can clearly not be attributed to major differences in writing expertise. L1 is the only factor that plays a role in this comparison, which is not to say that the observed differences are necessarily caused by transfer. Chapter 5 will consider if and to what extent the higher mean use of linking adverbials, as well as their steady increase between year 1 and 3 in the literature essays, might possibly be teaching-induced.

The one other persistent difference between the Dutch learners and the native speakers relates to the way they use their adverbials, regardless of adverbial

class, for local anchoring. Even by year 2 the PoEC for local anchors in argumentative essays written by Dutch learners and NSs is only 0.002, although if we had had any year 3 argumentative data for the learners, the PoEC might well have turned out higher. The PoECs for the NNS literature essays and NS literature articles, on the other hand, go from 0.000 in year 1, to 0.005 in year 2, to 0.082 in year 3. In spite of remaining differences, therefore, there is clearly a development in the direction of native writing. Apart from the fact that Dutch learners use more local anchors than the NS authors of the most comparable reference corpora, something else seems to be going on as well. The NS authors of the LOCNESS argumentative essays use considerably fewer local anchors than the NS authors of the VUNSPRAC articles, while the VUNSPRAC articles on linguistics and literature are remarkably similar. This might either be an effect of writing expertise or register, although the latter seems the most likely option as the Dutch learners also use more local anchors in their literature essays, which are more academic in purpose and tone, than in their argumentative essays, which are produced in the context of the writing curriculum and therefore focus on writing skills and argumentation rather than academic content. Judging from their decreasing use of adverbials used for local anchoring as well as different patterns of use in argumentative and literature essays, it would seem that Dutch learners' increasing awareness of cross-linguistic differences between Dutch and English in the way that local anchors are used for discourse linking goes hand in hand with a remarkable sensitivity to register in their L2.

Having established that there are considerable differences in the way Dutch EFL learners and NSs use PSAs for discourse linking, the next step is to establish what exactly causes this discrepancy. Following Jarvis' criteria for establishing transfer (2000), chapter 5 investigates Dutch EFL learners' use of PSAs in their L1, to assess the degree of L1-interlanguage congruity, while chapter 6

considers the interlanguage of advanced EFL learners with other L1 backgrounds, to determine to what extent overuse of circumstance adverbials, linking adverbials and local anchors might be a common feature of learner language.

Our observation that the relatively high frequency of PSAs is one of the subtle features that distinguishes advanced Dutch learner writing from NS writing does not mean to imply that learners should strive to achieve an ideal NS 'norm'. Our study does not aim to feed into a prescriptive native-speaker model but to offer a descriptive account of one of the defining features of Dutch EFL learners' 'discourse accent' (Scarcella, 1983) and of the nature of language development at advanced stages of acquisition. We would nevertheless like to argue that unlike other language users, (future) language professionals might benefit from an awareness of features of different varieties of NS and NNS writing and their potential communicative effects. Such awareness enhances learners' ability to develop into autonomous language professionals who have the linguistic finesse to write, translate, edit or teach communicatively appropriate and effective English.

5. Instant cohesion:

Exploring the role of transfer and teaching in the use of cohesive adverbials in L2 English writing¹³

Abstract

This paper investigates the role of transfer and teaching in a distinctive feature of Dutch learners' written English: frequent use of initial adverbials to achieve cohesion. A corpus analysis of L1 Dutch, L1 English and Dutch learners' L2 English reveals that the incidence of both cohesive and non-cohesive initial adverbials in L1 Dutch far exceeds their occurrence in L2 and L1 English. Only cohesive adverbials were also used more frequently in L2 English than L1 English. This partial L1-interlanguage congruity points to the role of transfer. However, a transfer-based account fails to explain why frequent use of non-cohesive adverbials in L1 Dutch is not mirrored in L2 English. Scrutiny of secondary-school English and Dutch coursebooks and preparatory material for the national final reading exam suggests that Dutch learners' transfer-induced use of cohesive adverbials is reinforced by a largely reductionist approach to teaching textual cohesion, leading learners to believe that repeated use of linking words will help them achieve instant cohesion.

Keywords: linking words, cohesion, transfer, teaching effect, L2 English

¹³ This is a revised version of: Van Vuuren, S. & De Vries, R. (forthcoming). Instant Cohesion: exploring the role of transfer and teaching in the use of cohesive adverbials in L2 English writing. In P. de Haan, R. de Vries & S. van Vuuren (Eds), *Language, Learners and Levels: Progression and Variation*. Louvain-la-Neuve: Presses universitaires de Louvain.

1. Introduction

Learning how to achieve cohesion is generally considered to be an important teaching goal in both L1 and L2 writing education. As students progress and start writing longer texts about more complex topics, the readability of these texts depends increasingly on the successful application of cohesive strategies such as reference, substitution, ellipsis, conjunction and lexical cohesion (Halliday & Hasan, 1976). The previous chapter has demonstrated that clause-initial adverbials which serve a cohesive purpose – both linking adverbials and adverbials of any adverbial class that are used for local anchoring – occur much more frequently in argumentative essays written by advanced Dutch learners of English than in those written by L1 writers of comparable age and educational background. This is perhaps not surprising, considering the fact that Dutch is a verb-second (V2) language, with a dedicated preverbal slot for discourse linking (cf. Verheijen *et al.*, 2013). At the same time, Dutch learners of English are clearly not exceptional in their overuse of (adverbial) connectors. Studies into the writing skills of Hong Kong, Taiwanese, Chinese and Polish learners have also reported overuse of linking devices¹⁴, for which the authors have proposed several intertwined and partially overlapping explanations. Crewe (1990), Field and Yip (1992), Milton and Tsang (1993) and Leńko-Szymańska (2008) have all argued that teaching and coursebooks might play a role. Contextless, atomistic listing in coursebooks might create the impression that non-equivalent linking devices are in fact interchangeable, that the more learners use them, the more structured and coherent their text will be. In other words, that instant cohesion can be achieved

¹⁴ What we here refer to as linking devices (a.k.a. connectors or cohesive devices) is in most studies understood to include both linking adverbials and conjunctions, in any position in the sentence, although some focus on linking adverbials in particular.

by liberal use of linking devices. This may lead to what Crewe refers to as 'surface logicality' (1990).

L2 learners' preferences for placement of linking adverbials within the sentence also need to be taken into account. While Granger and Tyson (1996) and Altenberg and Tapper (1998) did not find evidence of overuse of linking adverbials among French or Swedish learners of English, they observed that both learner groups place these adverbials in clause-initial position more frequently than native speakers do. Gilquin *et al.* (2007) also note that learners with various L1 backgrounds commonly place 'however' in clause-initial position, while native speakers tend to place them in medial position relatively more often. They suggest that if this preference for initial position applies to other linking adverbials as well, it may well be an interlanguage universal.

The present study aims to examine the roles of transfer and teaching in the persistent overuse of clause-initial adverbials in English argumentative essays written by advanced Dutch learners of English. The first part of this paper investigates to what extent Dutch EFL learners have the same preference for clause-initial adverbials in their L1 and their L2, which would substantiate a transfer-based account. To this end, we compared the longitudinal L2 English corpus and L1 English corpus described in the previous chapter with a semi-longitudinal corpus of L1 Dutch argumentative essays, written by three cohorts of Dutch university students of English Language and Culture in their first, second and third year. The second part of this chapter explores the role of L1 Dutch and L2 English secondary-school coursebooks and final secondary-school reading exams in Dutch learners' overuse of clause-initial adverbials. We will consider the frequency with which different types of adverbials are discussed as well as the emphasis placed on text structuring in general. We will also look into the way

coursebooks present these adverbials: whether or not context is provided and attention is paid to functional, grammatical and register differences.

2. Methodology

2.1 Corpus construction and annotation

We compared the L1 English novice writing included in LOCNESS (Granger, 1996) with the L1 Dutch and L2 English novice writing in the Longitudinal Database of Learner English (Meunier, 2015), produced by L1 Dutch university students of English Language and Culture. In order to guarantee comparability between the different (sub)corpora, we included timed argumentative essays only, which unfortunately meant we could not include any year 3 English texts written by the Dutch learners (Table 1).

Table 1. Corpus composition

	Language	texts	words	mean	SD
LONGDALE year 1	English	92	25,532	277.52	113.59
LONGDALE year 1	Dutch	45	17,816	395.91	114.69
LONGDALE year 2	English	77	22,298	289.58	93.67
LONGDALE year 2	Dutch	47	18,096	385.02	141.82
LONGDALE year 3	Dutch	56	22,779	406.77	169.78
LOCNESS	English	110	59,229	538.45	181.79

Syntactic annotation was added to the English and Dutch texts by means of the Stanford Parser (Klein & Manning, 2003) and the Alpino Parser (Van Noord, 2006) respectively. Clause-initial adverbials in declarative main clauses were subsequently filtered out using Corpus Studio (Komen, 2012). The output of this Corpus Studio query was a database of 888 clause-initial adverbials for the L1 Dutch texts, 665 for the L2 English texts and 558 for the LOCNESS A-levels.

Annotation was then added in the Cesax corpus-annotation tool (Komen, 2012), with each adverbial being categorised according to adverbial class and referential state (cf. chapter 1). We distinguished between (1) circumstance adverbials, which provide circumstantial information about the proposition, such as ‘in this essay’; (2) stance adverbials, which are used to comment on the content or style of the message, such as ‘personally’; (3) linking adverbials, which serve to create cohesion, such as ‘however’ (Biber *et al.*, 1999). For each of these adverbials we also noted whether or not they were used for local anchoring.

2.2 Statistical analysis

Following the study reported on in chapter 4, the corpus data were statistically analysed by adopting a Bayesian perspective (cf. Littré, 2014).

Whereas traditional frequentist statistics is concerned with the question whether the null hypothesis that two distributions do not differ needs to be rejected under a given significance level α , usually set at .05, this Bayesian approach results in a posterior distribution which relates each possible outcome to the probability of this outcome being the true value of the coefficient, given the data.

In order to facilitate the analysis of the posterior distributions, two additional scores were computed: (1) the mean of the posterior distribution, which takes both the most likely value into account, as well as the uncertainty (or variance) of the distribution, by weighing each value on the x-axis by its corresponding probability density, and (2) the Probability of Equal Coefficients (PoEC), which indicates the probability that the parameter values for a set of two groups are identical. For a more detailed description of the statistical procedure, the reader is referred to chapter 4.

3. Corpus findings

The posterior distributions of the L1 English, L2 English and L1 Dutch essays show that the Dutch learners use considerably more clause-initial adverbials in their L2 English texts than the L1 English writers do, with almost completely non-overlapping distributions both in year 1 and in year 2 (Figures 1 and 2) and a 0.0000 PoEC (Table 3). Still, both Figures 1 and 2 and PoECs listed in Table 3 demonstrate that the gap between what the Dutch EFL learners do in their L1 and in their L2 is even more pronounced. The frequency with which Dutch novice writers use clause-initial adverbials in their L1 far exceeds both the frequency found in their L2 English writing and in L1 English novice writing.

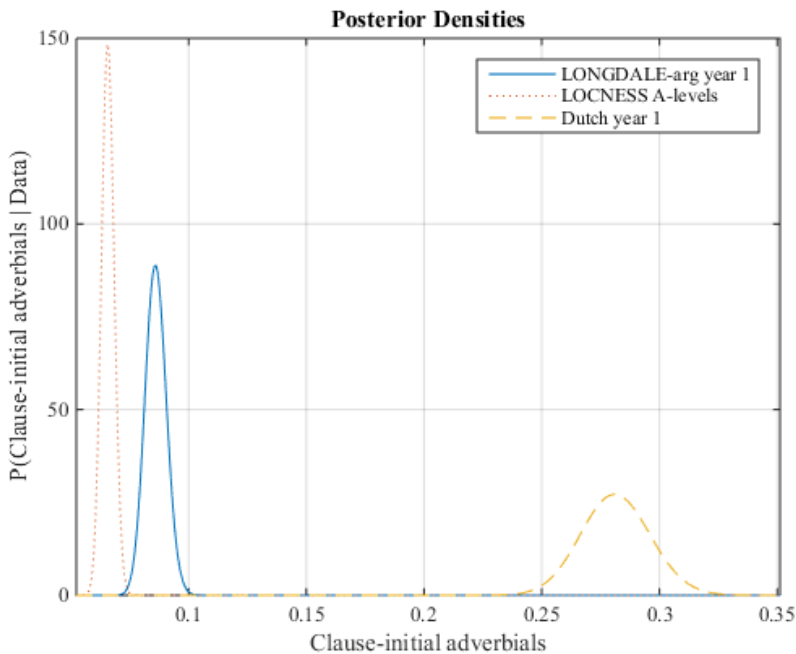


Figure 1. Initial adverbials in L1 English, L2 English and L1 Dutch (year 1)

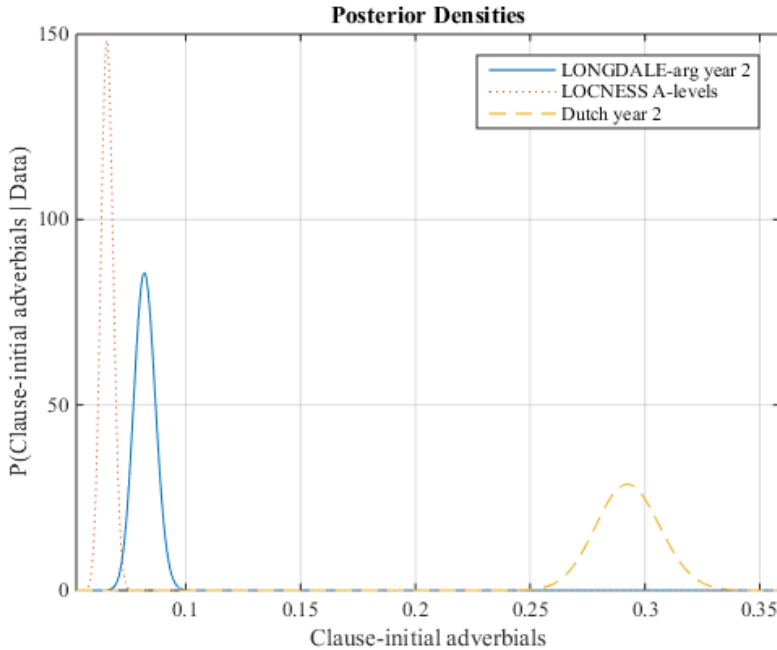


Figure 2. Initial adverbials in L1 English, L2 English and L1 Dutch (year 2)

The higher mean use of clause-initial adverbials in L1 Dutch writing compared with L1 English and L2 English writing is reflected in a considerably higher mean use of each of the three main adverbial classes – circumstance, stance and linking – as well as of those adverbials that are used for local anchoring (Table 2). The gap between L1 Dutch texts and both L1 and L2 English texts is further confirmed by the fact that for all adverbial categories PoECs do not rise above 0.0000, neither in year 1 nor in year 2 (Table 3).

Table 2. Mean values initial adverbials in L1 English, L2 English and L1 Dutch

	L1 English	L2 English		L1 Dutch		
		Yr1	Yr2	Yr1	Yr2	Yr3
Initial adverbials	0.0657	0.0861	0.0822	0.2814	0.2927	0.2949
Circumstance	0.0270	0.0336	0.0274	0.1311	0.1735	0.1180
Stance	0.0124	0.0092	0.0147	0.0480	0.0385	0.0449
Linking	0.0266	0.0438	0.0407	0.1045	0.0826	0.1339
Local anchors	0.0022	0.0092	0.0066	0.0512	0.0629	0.0749

Table 3. PoECs initial adverbials between English, L2 English and L1 Dutch

	L1 English/ L2 English		L1 Dutch/ L2 English		L1 Dutch/ L1 English	
	Yr1	Yr2	Yr1	Yr2	Yr1	Yr2
	Initial adverbials	0.0000	0.0005	0.0000	0.0000	0.0000
Circumstance	0.0175	0.1206	0.0000	0.0000	0.0000	0.0000
Stance	0.0559	0.1085	0.0000	0.0000	0.0000	0.0000
Linking	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Local anchors	0.0000	0.0016	0.0000	0.0000	0.0000	0.0000

The Dutch novice writers use more linking adverbials and local anchors in their L2, English, than the L1 English novice writers do, again with non-overlapping distributions and 0.0000 PoECs in both cases (Figures 3 and 4 and Table 3).

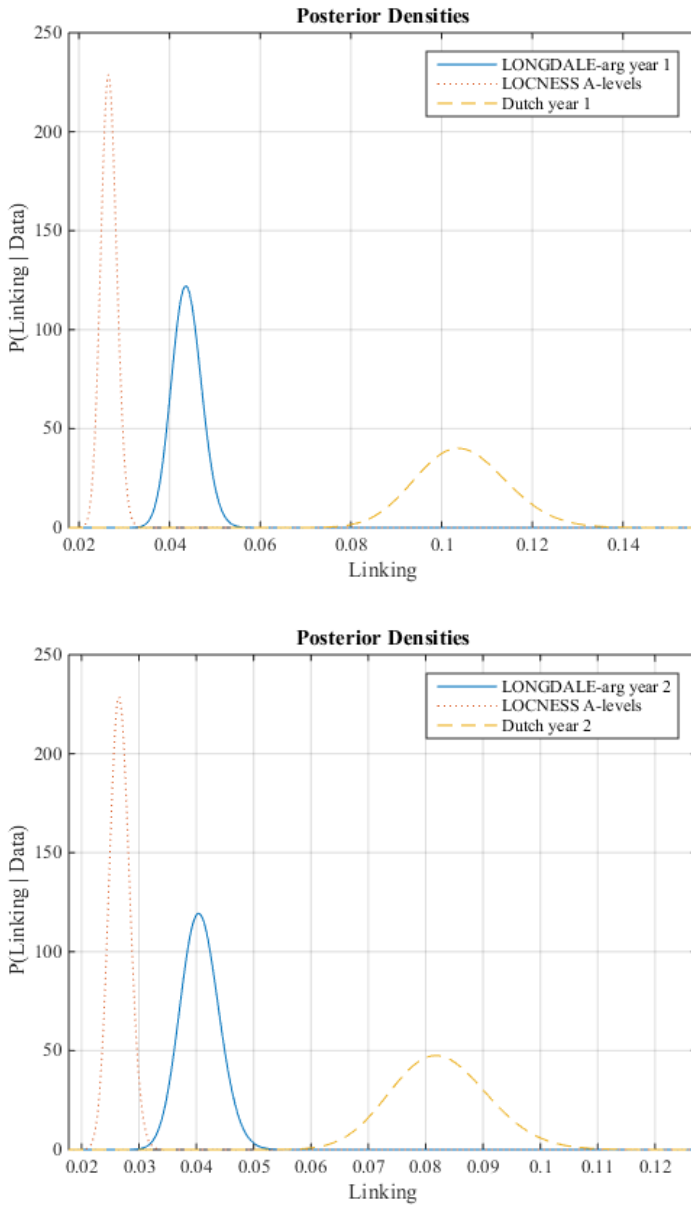


Figure 3. Initial linking adverbials in L1 English, L2 English and L1 Dutch

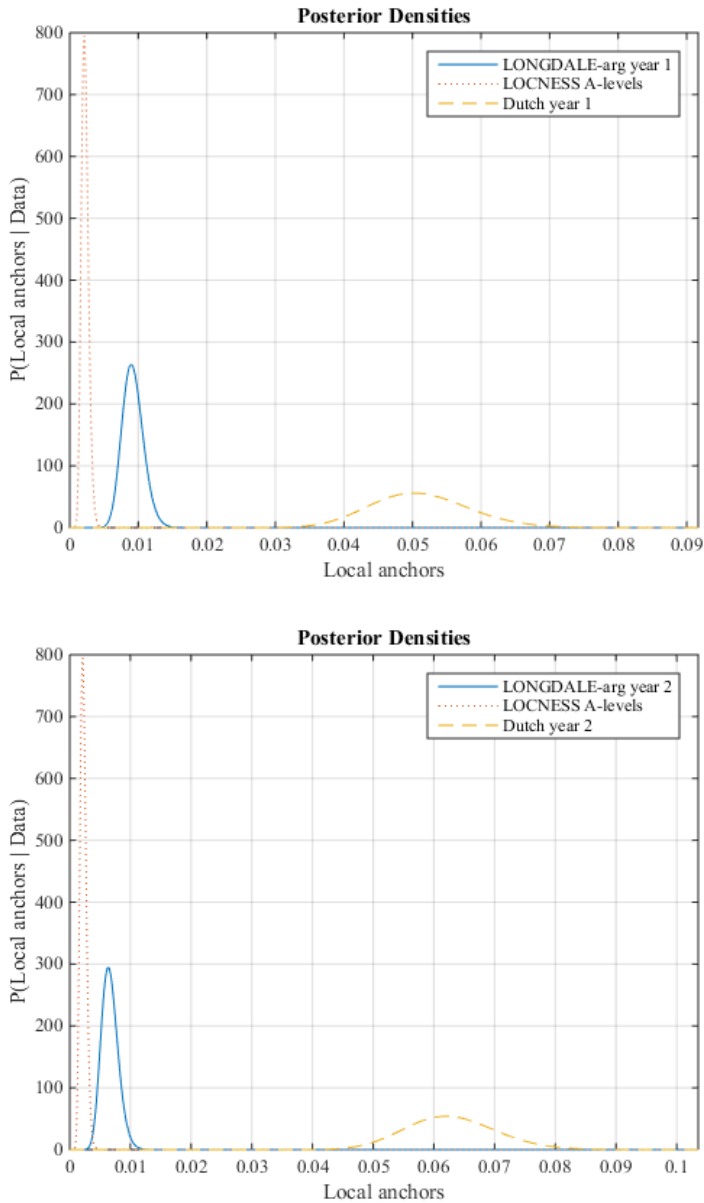


Figure 4. Initial local anchors in L1 English, L2 English and L1 Dutch

The difference between L1 and L2 English essays in the use of circumstance and stance adverbials, however, is less evident. Dutch learners initially use slightly

more circumstance adverbials in their English essays than the L1 English novice writers do, but by year two the mean values of both groups are very similar and the PoEC has risen to 0,1206. Likewise, it is difficult to maintain there is a reliable distinction between L1 and L2 English essays in the use of stance adverbials. In spite of the fact that the Dutch novice writers use more stance adverbials in their L1 than the L1 English novice writers do, this apparently does not translate directly into an increased use of stance adverbials in their L2. In fact, in their English essays they initially use fewer than the L1 English novice writers do. By year 2 they have started to use slightly more, but at 0,1085 the PoEC between L1 and L2 English essays suggests there is no reliable difference between both groups.

All in all, there does seem to be some evidence for one of Jarvis' (2000) criteria for establishing transfer: L1-interlanguage congruity, but only for the class of linking adverbials and those adverbials that are used for local anchoring. Transfer might therefore go some way towards explaining Dutch learners' natural tendency to achieve cohesion by placing adverbials in clause-initial position. But transfer falls short of explaining why linking adverbials as well as local anchors should be subject to transfer from Dutch, whereas circumstance and stance adverbials are not. Stance adverbials in particular stand out because Dutch learners initially use them less frequently than L1 English writers do, whereas in their L1 they use considerably more. In view of this not entirely satisfactory transfer explanation we will continue by exploring to what extent secondary-school coursebooks as well as final reading exams for both Dutch and English enhance Dutch learners' tendency to use initial adverbials for discourse linking.

4. The role of teaching

4.1 English coursebooks in Dutch schools

Lower-level coursebooks, such as those at *CEFR* A1-A2, hardly pay any attention to coherence and cohesion, as they mostly focus on spoken language. Any writing assignments at this level tend to be very short and what few linking devices are discussed, notably ‘and’, ‘or’, ‘but’ and ‘because’, tend to operate at sentence level. This is in line with the *CEFR* scale for coherence and cohesion at A1 and A2 (Council of Europe, 2001, p. 125). We therefore decided to concentrate on the three most commonly used titles at *CEFR* B1 and higher, used in years 4, 5 and 6 of pre-university education (VWO):

1. Noordhoff's *Stepping Stones*: the most frequently used title in Dutch secondary education (according to publisher) - specifically aimed at Dutch learners of English
2. OUP's *New Headway*: worldwide number one coursebook (according to publisher), also popular in the Netherlands

Many schools in the Netherlands offer learners an opportunity to take a Cambridge English exam and many of our students turn out to have taken that opportunity.¹⁵ We therefore also looked at one of the more popular titles to prepare for these Cambridge English exams:

3. Pearson's *Gold*: commonly used in Dutch schools as preparation for Cambridge English exams

¹⁵ 27% of students in cohort 2013 have taken either a Cambridge Advanced English or Cambridge Proficiency in English exam.

Scrutiny of each of the coursebooks revealed that there are noticeable differences in the way linking devices are treated and presented.

Occurrence of linking devices in coursebooks

We categorised all adverbials that occurred in exercises or reference sections in each of the coursebooks according to Biber's (1999) classification, distinguishing between circumstance, stance and linking adverbials. As adverbials tend to be presented in the context of writing assignments, the attention they receive is commensurate with the number of writing assignments offered. *Stepping Stones* pays relatively little attention to writing practice. It is probably for this reason that adverbials feature less prominently in *Stepping Stones* in comparison with *Gold* and *New Headway*. By extension, the type of writing exercise also appears to determine what categories of adverbials students are presented with. The more narrative-oriented writing activities offered in *New Headway* go a long way towards explaining the predominance of circumstance adverbials, mostly time adverbials, which help to structure a narrative. This genre is not part of the Cambridge advanced and proficiency writing exams and it is hardly surprising, therefore, that *Gold* presents far fewer circumstance adverbials. Stance adverbials are presented in *New Headway* as discourse markers that can be used either in personal writing assignments, such as informal letters, or in conversation. In *Gold* stance adverbials feature mostly in the context of preparation for the review that is part of the writing paper. Unlike circumstance and stance adverbials, which cannot be captured in a finite set, linking adverbials lend themselves very well to presentation in atomistic lists. The frequencies of linking adverbials in *New Headway*, *Gold* and *Stepping Stones* are therefore mainly determined by their occurrence in reference sections. An overview of all circumstance, stance and linking adverbials in each of the coursebooks is provided in Appendix I.

Presentation with or without context

Linking devices are often presented in lists, primarily in reference sections, which are divided into semantic categories, such as contrast, result, reason, purpose and addition. *Gold* and *New Headway* provide one-sentence contexts for each of these linking devices, whereas *Stepping Stones* does not provide any context or definition at all. This practice of atomistic listing leads to linking devices such as ‘on the contrary/on the other hand’ or ‘therefore/thus’ being presented as interchangeable alternatives in *Stepping Stones* (Hartog *et al.*, 2009). While atomistic listing does not in itself lead to overuse, it may lead learners to assume that they can achieve cohesion by picking, choosing and using linking devices.

Instructions for usage

This assumption is reinforced by repeated advice to learners to use linking adverbials to structure their writing. For example, *Stepping Stones* warns learners to ‘[m]ake sure you use the Language help on linking words to make your story coherent’ (Hartog *et al.*, 2009, p. 22). This implies that use of cohesive strategies equals coherence and, by extension, absence results in incoherent text. The frequency with which such instructions are offered varies considerably. Because the *Gold* series prepares learners for a Cambridge English exam, in which they are assessed on their ability to use ‘a range of linking words effectively’ (Thomas & Burgess, 2014, p. 206), it is only logical that the authors should repeatedly emphasise the importance of using linking words to pass the writing paper. This is done in boxes with ‘Exam Strategies’ and ‘Exam Tips’ that feature prominently in the book:

Language Tip

Linking adverbials connect one sentence in a logical way to another. They usually appear at the beginning of a sentence and are followed by a coma. They help to make writing more cohesive.

(Thomas & Burgess, 2014, p. 137)

Exam Tip

Remember to use linking adverbials to organise your ideas and to make it easy for the reader to follow your argument

(Thomas & Burgess, 2014, p. 138)

Against this backdrop of repeated encouragement to structure texts by means of linking devices, *Gold Proficiency* is the only coursebook that on two occasions warns learners that linking devices ‘should only be used when necessary, not in every sentence’ (Newbrook & Wilson, 2003, p. 180) and that overuse is ‘unnatural’ (p. 232).

Exercises

There is a wide variety of exercises in which the use of adverbials is practised. These exercises mostly focus on the use of linking adverbials, or linking devices in general. An overview of the various types of exercises is presented in Table 4.

Table 4. Numbers of different types of exercises on adverbials in *New Headway*, *Gold* and *Stepping Stones*

	<i>New Headway Upper- Intermediate</i>	<i>New Headway Advanced</i>	<i>Gold Advanced</i>	<i>Gold Proficiency</i>	<i>Stepping Stones VWO 4</i>	<i>Stepping Stones VWO 5</i>	<i>Stepping Stones VWO 6</i>
Rewrite		2		2			
Connect sentences		2		3			
Find in text/tapescript	3	2	2	1			
Fill the gaps	4	6	3	1			
Insert in correct position	1						
Use in writing (or speaking) assignment	1	4	5	3	2	1	1
Complete sentence starting with linking word	1			1			
Register awareness				1			
Match adverbial to semantic function			2	1	1		

What stands out is the scarcity of exercises in *Stepping Stones*. On a number of occasions learners are asked to use adverbials in speaking and writing assignments and once they are asked to match adverbials to their semantic function. Learners are never asked, however, to familiarise themselves with the use of these adverbials in context. *Gold* and *New Headway* take a more integrated approach, offering rewriting, gap-fill and awareness-raising activities, next to the writing and speaking assignments in which learners are encouraged to apply what they have learned. The input for these awareness-raising tasks seems to have a higher

incidence of linking devices than might be the case in truly authentic materials. The frequency and choice of linking devices in the following informal exchange in *New Headway*, for example, might be slightly unnatural:

N: I want to do something different for our holidays this year. **Ideally**, something adventurous for a change.

A: Really, what sort of thing.

N: Dunno, actually. **Maybe** going to a lake and learning watersports. **Apparently**, windsurfing is exciting and we could also learn to sail.

A: It doesn't sound much fun to me. **Personally**, I prefer something a bit more relaxing. **Basically**, I think we work hard enough all year so that we deserve to do nothing somewhere nice for a couple of weeks.

N: But I'm fed up with lying on a beach and all that. We can do that any old time. It's high time we had some new experiences in life. **After all**, we don't get much holiday a year.

A: **Well, obviously**, I'm very impressed with your new lease of life. **However**, I'm still not sure what you have in mind or if I want to do it, too.

[...]

(Soars, Soars & Wheeldon, 2005, p. 85)

After listening to this conversation, learners do a gapfill exercise focusing on a number of these 'linking and commenting expressions', after which they are invited to identify even more such expressions in the tapescript. It is unavoidable that coursebook writers incorporate more linking devices in their texts than one would expect to find in authentic, undoctored texts. After all, the learners' attention needs to be drawn to the use of these devices in context. On the

downside, learners might get the wrong idea about both the frequency and register appropriacy of linking devices in natural language.

Alternative cohesive strategies

Obviously, cohesion is not exclusively or even predominantly dependent on the use of linking devices. Apart from linking adverbials and conjunctions, *New Headway* also discusses categories of circumstance and stance adverbials that can be used to achieve cohesion, as well as participles and infinitives which serve to connect ideas on sentence level. *Gold* is even more explicit in offering alternative means of achieving lexical and grammatical cohesion, i.e. reference, substitution, ellipsis and conjunction (Thomas & Burgess, 2014, p. 142, p. 171). Alternatives like these are only hinted at in *Stepping Stones*, when they say that in order to achieve coherence ‘we use linking words/connectors, among other things’ (Hartog *et al.*, 2008, p. 31). The ‘other things’ remain unspecified.

All in all, the frequency with and manner in which linking devices are presented and practised in *Stepping Stones*, *New Headway* and *Gold* is not consistent enough to fully explain our students’ use of initial cohesive adverbials. The fact that the *Gold* series prepares students for a Cambridge English exam, which has a substantial writing component, is reflected in its focus on cohesion. It also tests students at a higher level (C1/C2) than the pre-university final exams (B2) which *New Headway* and *Stepping Stones* prepare students for. This might explain why *Gold* pays more attention to alternative means of achieving cohesion and also offers stylistic advice, even warning students not to overuse linking adverbials. We will now consider if and to what extent the final pre-university exams, which are the culmination of the secondary-school English curriculum, may encourage learners’ focus on linking devices.

4.2 Final secondary-school reading exams

Linking devices also feature heavily in the context of preparation for the nationwide reading exam as set by the central institute for test development (CITO). Interviews with four English teachers at schools previously attended by students of English Language and Culture at Radboud University Nijmegen confirmed a widely-held belief that knowledge of linking devices is of vital importance for passing the exam. Teachers report that exam training starts as early as year 1. Exam preparation tends to focus heavily on top-down reading strategies, such as predicting content from headlines and pictures, activating knowledge about the subject, making predictions on the basis of the first and last sentence of each paragraph, and focusing on linking devices.

The importance of linking devices is further underlined by the widely-used annual *Examenbundel* (Van Putten & Verploegh, 2015) and *Examentraining* (Honders *et al.*, 2014), both of which offer a collection of previous exams with vocabulary lists, exam strategies, etc. Both offer a list of what they either call ‘signalling words’ or ‘structuring words’, including adverbs as well as coordinating and subordinating conjunctions. Whereas the *Examenbundel* offers these without context, but with an indication of the logical relation they signal, the *Examentraining* provides a short sample sentence for each ‘structuring word’. The *Examenbundel* encourages students to study the list of linking devices and even learn them by heart, in order to facilitate the interpretation of logical relations and warns learners to ‘[p]ay attention to them, because signalling words are an important tool in answering reading exam questions’ (Van Putten & Verploegh, 2015, p. 84, our translation). It also mentions a lack of knowledge of vocabulary in general, and linking devices in particular, as a common reason for failing the exam (6). Similarly, *Examentraining* points out that ‘structuring words’ are of great

importance in both traditional multiple-choice questions and multiple-choice gapfills and recommends learning them by heart (Honders *et al.*, 2014, p. 299).

Pre-university reading exams include a handful of questions that directly test knowledge of linking devices. The complete set of 20 final reading exams that were administered nationwide between 2006 and 2015 contain a total of 831 questions. Not even 7% (56) of these questions require learners to choose a connector either to fill a gap or as an alternative beginning for a sentence or paragraph. Knowledge of linking devices used in the exam texts themselves may in a number of cases indirectly contribute to learners' ability to choose the correct answer to a multiple-choice question, although this is difficult to quantify. 3% (26) of all questions test learners' ability to recognise the function of a particular paragraph or the (logical) relation between two paragraphs. The interpretation of these relations might conceivably be facilitated by recognition of linking devices, but in most cases these do not feature in the paragraphs in question.

Not all of the linking devices of the recommended 'list of signalling words' occur particularly frequently in the reading exams themselves, with 22% of linking devices (e.g. 'all the same', 'as a matter of fact', 'in brief', 'in conclusion', 'lastly') not appearing either in the texts or the questions and another 12% appearing only once or twice in ten years of final reading exams. Even the linking devices that do occur more than once per exam on average are not always picked up on in questions. Examples are 'actually' (27 times), 'as well as' (22 times), 'rather than' (38 times) and 'probably' (22 times). It is also interesting to note that the linking devices that are offered as alternatives in multiple-choice questions do not necessarily reflect the type of language used in the texts themselves. The word 'nevertheless', for example, is used no fewer than 15 times in exam questions, but occurs only 4 times in 106,042 words of exam texts. The same applies to

‘paradoxically’, which occurs 5 times in exam questions and not at all in the texts themselves.

Apart from their use in answering exam questions, focusing on linking devices might be considered an important reading strategy to make up for a lack of vocabulary knowledge. Texts are reputed to contain so much low-frequency vocabulary that learners would do well to hang on to the logical relations expressed by linking devices. In order to assess quite how much low-frequency vocabulary the exam texts contain, we processed all exam texts with the *Compleat Lexical Tutor* (Cobb). This produced a vocabulary profile based on the combined British National Corpus (BNC) and Corpus of Contemporary American English (COCA) word frequency lists. Analysis of our corpus shows that 77.19% of words in the texts can be found in the 1,000 word list and another 10.15% in the 2,000 word list, with 6.53% of the texts made up of words from the 3,000 word list and 6.13% of less frequent vocabulary. Estimates of the percentage of words in a text learners need to know in order to achieve adequate comprehension vary between 95 and 99% (Schmitt *et al.*, 2011, p. 26). It seems unlikely that learners at pre-university level are familiar with 95% of the vocabulary used. Administration of the Nation 3,000 and 5,000 vocabulary tests (Laufer & Nation, 1999) showed that the number of our first-year students of English Language and Culture that manages to pass the test’s threshold of 80% correct answers (shaded) for the 3,000 and 5,000 word lists is 75.76% and 24.4% respectively (Table 5).

Table 5. First-year students' results for Nation 3,000 and 5,000 vocabulary tests

% of correct answers	Nation 3,000	Nation 5,000
100%	7 (17.1%)	1 (2.4%)
94%	4 (9.8%)	2 (4.9%)
88%	7 (17.1%)	3 (7.3%)
83%	13 (31.76%)	4 (9.8%)
77%	1 (2.4%)	3 (7.3%)
72%	5 (12.2%)	5 (12.2%)
66%	3 (7.3%)	3 (7.3%)
61%	1 (2.4%)	7 (17.1%)
55%	0 (0%)	6 (14.6%)
< 50%	0 (0%)	7 (17.1%)
	41 (100%)	41 (100%)

Of course, the percentage of known vocabulary required also depends on the degree to which a text needs to be understood. Being able to answer exam questions does not necessarily presuppose full understanding of the text. Schmitt *et al.* (2011) also suggest that other variables play a role, including grammatical knowledge, L1 reading ability, inferencing ability and awareness of discourse structure (p. 29). In this context it is interesting to note that research into the role of text and question variables in multiple-choice L2 Dutch exams (Hoek, 2012) revealed that the difficulty of these exams can be predicted mostly on the basis of question variables, such as lexical overlap of the correct answer with the text and the plausibility of distractors. Lexical or syntactic complexity of the text itself did not significantly predict L2 Dutch exam results. It is not beyond the realm of imagination that this applies to L2 English reading exams as well, although this remains to be confirmed by future research.

4.3 Teaching transfer

In order to come to a better understanding of students' motivation for their frequent use of linking devices, we administered a questionnaire to a cross-section of our students. They were asked how important they feel the use of linking devices is in general and to what extent linking devices were discussed in their secondary school English and Dutch classes. Over 80% of all 53 respondents considered the use of linking words to be one of the three most useful strategies in producing a reader-friendly piece of text. Interestingly, while roughly half of the respondents indicate that they were encouraged to use linking devices in their pre-university English classes, the other half say that they were not discussed, including a few reporting not being able to remember. More respondents pointed to the role of their Dutch classes, with 70% of respondents reporting that the use of linking devices was actively encouraged. Respondents indicate that they had to learn lists of linking words and the logical relations that they signal, and that they were considered to be crucial in the context of both writing and reading exams. Some even reported being awarded extra points for each linking device used or being punished for not using enough.

This impression is borne out by scrutiny of *Nieuw Nederlands*, *Talent* and *Op Niveau*, the three most commonly-used coursebooks for L1 Dutch in pre-university education (Meestringa & Ravesloot, 2013, p. 7). Compared with the English coursebooks discussed above, L1 Dutch coursebooks attach considerably more importance to linking devices, which they commonly refer to as 'signalling words'. All three coursebooks offer lists of linking devices and the logical relations they express as well as practice in the form of awareness-raising tasks and occasional gapfill exercises. Students are also advised to use linking devices to structure a wide variety of writing assignments such as expository and argumentative texts as well as summaries:

Write your summary in well-structured sentences that are joined together. Therefore, make sure you use signalling words, referencing words and repetition.

(Jongsma *et al.*, 2010, p. 63, our translation)

Unlike L2 English coursebooks, which hardly pay attention to linking devices until the final years of pre-university education, L1 Dutch coursebooks emphasise the necessity of using linking devices to signpost logical relations from year one onwards. This may not be surprising, considering the difference in proficiency between learners' use of English as an L2 and Dutch as an L1. Not only are linking devices discussed at an earlier point, they are also presented with greater frequency and consistency than in the L2 English coursebooks. Linking devices in Dutch coursebooks are predominantly discussed in the context of text analysis exercises, which serve as preparation for the Dutch final secondary-school reading exam. Learners are repeatedly directed to linking devices in the text in order to facilitate recognition of logical relations within and between paragraphs. Exercises require learners to find linking devices in a text and indicate what logical relation they signal, or supply a linking word to signpost a logical relation that was left unmarked, as in the following example from *Talent*:

- a) What causal signalling words can you find in the sentences? There are four.
- b) Which sentence does not contain a signalling word?
- c) What signalling word could you use in that sentence, so that opinion and argumentation are presented more clearly?

(Mulder *et al.*, 2014, p. 184, our translation)

It seems likely that after six years of training, learners are imbued with a strong sense of the importance of linking devices for text structuring. This might stand them in good stead because scrutiny of the Dutch final reading exams between 2006 and 2015 revealed that they focus heavily on the identification of logical relations between sentences and paragraphs. In contrast to the final reading exams for English, no questions require the candidate to demonstrate their knowledge of discrete linking devices by filling the gaps in a text. Instead, linking devices are used as a tool to guide learners in answering questions on structure and logical relations. For example, when learners are asked to identify a number of causes, they need to see that the relevant sentences in the text are introduced by phrases such as '*allereerst*' (first of all) and '*ook*' (also). Similarly, when learners are asked to determine the argumentative function of a paragraph in relation to a preceding paragraph, attention to linking words is likely to steer them in the right direction. Although it is difficult to quantify the role of linking devices in questions, roughly 38% of exam questions explicitly test learners' understanding of structure, logical relations or argumentative functions, for which they might need to rely on linking devices. We are therefore inclined to believe that in order for learners to pass the L1 Dutch reading exam, teachers and learners are indeed well-advised to focus on the role of linking devices in text structuring. This focus goes some way towards explaining why learners arrive at university believing that texts hang together by virtue of linking devices.

5. Discussion and conclusion

Dutch university students of English Language and Culture have been shown to use more initial linking adverbials and local anchors in their L2 English texts than L1 English writers. This paper set out to assess the extent to which transfer from Dutch and/or teaching practices might contribute to this feature of Dutch L2

English writing. We found that L1 English texts contain considerably fewer initial linking adverbials and local anchors not only in comparison with L2 English writing but also with L1 Dutch writing. This L1-interlanguage congruity suggests that transfer from Dutch might to some extent be responsible for the frequency with which Dutch learners use initial linking adverbials and local anchors. A transfer-based account fails to explain, however, why the exceptionally frequent use of stance and circumstance adverbials in L1 Dutch should not be transferred to the same degree as linking adverbials and local anchors, both of which serve a cohesive purpose.

It would appear that in coursebooks cohesion is often reduced to the use of linking devices, which is perhaps more manageable than other strategies, both from a teaching and a learning perspective. After all, they can be presented in neat lists, subdivided into different logical relations, such as cause and effect, contrast, concession, exemplification etc. This might explain why more frequent use of circumstance, stance and linking adverbials in L1 Dutch writing compared with L1 English only leads Dutch learners to use more linking adverbials and local anchors, i.e. adverbials that serve a cohesive purpose, in their L2 English texts.

Scrutiny of three of the most frequently used L2 English coursebooks revealed that their presentation and treatment of linking devices differs considerably. *Stepping Stones* pays some attention to linking devices, but they are offered in contextless lists only. Both *Headway* and *Gold* emphasise the importance of linking devices in signposting logical relations, with *Gold Proficiency*, which prepares for an exam at *CEFR C2*, going one step further than the other coursebooks in that it also explicitly discusses alternative ways of achieving cohesion, pays some attention to register differences in the use of linking devices, and even occasionally warns against overuse.

The three L1 Dutch coursebooks we investigated pay even more attention to linking devices and also discuss alternative ways of achieving cohesion, e.g. repetition of key words, synonymy, pronouns and pronominal adverbs. However, they never warn against using too many linking devices. On the contrary, learners are advised to use them to flag each new argument, each counterargument, and each addition (Barends *et al.*, 2009, p. 33). Reading practice takes the form of text analysis exercises, in which linking devices are used as a tool to determine the logical relation between sentences or groups of sentences.

This heavy focus on linking devices seems to be warranted by the Dutch final reading exam, nearly half of which tests learners' understanding of structure, logical relations and argumentative functions. The English final reading exam has a different approach. Whenever linking devices are tested explicitly, they feature almost exclusively as options in gap-fill questions. The bulk of questions relate to content rather than structure. Nevertheless, linking devices feature heavily in dedicated exam preparation such as the *Examenbundel* and *Examentraining*. Learners might be felt to benefit from a focus on linking devices to keep track of the argumentation, especially if the presence of low-frequency vocabulary hampers understanding.

On balance, Dutch learners' natural tendency to use clause-initial position for discourse linking, which may be a transfer effect from Dutch, is likely to be reinforced by the treatment of linking devices in the English and Dutch secondary-school curriculum and final exams. It is difficult to assess the relative weight of each of these factors, as they appear to be inextricably intertwined. While the frequency of initial adverbials in L1 Dutch writing may point towards the role of transfer, it may just as easily be the direct result of the importance attached to linking devices in the L1 Dutch curriculum. This may in turn be a natural reflection of the fact that Dutch, as a V2 language, can easily accommodate

different types of constituents, including adverbials, in the preverbal slot and tends to use these for discourse linking purposes. A comparison of L1 Dutch novice and expert writing in relation to the frequency of initial adverbial use found in advanced Dutch learners' L2 English writing may shed further light on the respective roles of teaching and transfer.

The question is whether or not it matters that Dutch learners have a preference for achieving cohesion in their L2 English by means of initial adverbials. The communicative success of the average language user's writing is of course hardly dependent on a native-like use of adverbials. For the average language user communicative success may be the only relevant concern, but the learning objectives of students of English Language and Culture are likely to go beyond this. These students are training to become English language professionals, and may therefore benefit from an awareness of differences between Dutch and Anglo-Saxon writing culture, including ways of achieving cohesion and highlighting different elements in the sentence. This would allow them to produce what Ammon (2007) refers to as 'linguistically more refined texts' (qtd. in Flowerdew, 2015). Flowerdew (2015) concludes that '[i]f someone has chosen English as their field of professional endeavour, then it is only reasonable to expect them to have mastered this field to the highest level.' This does not automatically imply that an NS norm should be imposed on these learners, but it seems inevitable that learners should adopt some sort of standard, whether this standard should go under the name of 'native', 'natural', or 'general' English. Which model learners choose ultimately depends on their communicative needs, aptitude and motivation. It might also be argued that the frequency of cohesive adverbials is not simply a matter of language, let alone a matter of adherence to a near-native model. It is in some cases also a symptom of surface logicity (Crewe, 1990). If learners use logical connectors to achieve 'instant cohesion' that serves as a

shortcut for well-developed argumentation, this is likely to affect the readability of their writing and the credibility of their thinking.

6. Same difference?

L1 influence in the use of initial adverbials in English novice writing¹⁶

Abstract

The previous chapters have identified frequent use of clause-initial adverbials – linking adverbials and local anchors in particular – as characteristic of Dutch EFL writing. The present chapter presents a contrastive corpus analysis of novice writing by Dutch and Francophone learners as well as native speakers with the aim of determining whether this use of initial adverbials is a) a V2 transfer effect, b) a general interlanguage feature, independent of learners' L1, or c) a characteristic of novice writing in general, holding true for both native and non-native writers. We will show that both learner groups are 'equally different' from the native-speaker novice writers in their high frequency of initial adverbials, but appear to have distinct underlying reasons for this linguistic behaviour: Francophone writers place adverbials in initial position more often for stylistic purposes, while Dutch writers have a stronger tendency to use initial adverbials for local discourse linking.

Keywords: Contrastive Interlanguage Analysis, novice writing, EFL, transfer, pragmatics

¹⁶ This is a revised version of: Van Vuuren, S. & Berns, J. (submitted). Same Difference? L1 influence in the use of initial adverbials in English novice writing.

1. Introduction

Sentence structure and pragmatics interact. Lambrecht (1994, 16-17), for instance, argues that marked sentence structures in a given language signal the presence of some pragmatic feature, while the unmarked structure entails a pragmatically neutral reading. For second language learners, the pragmatics encoded in the linguistic structure of the target language seems to be a delicate aspect to master (cf. Dimroth & Starren, 2003; Callies, 2009). The way advanced Dutch learners of English as a Foreign Language (EFL) use pre-subject adverbials for discourse linking is a case in point. A Subject-Verb-Object (SVO) language like English and a Verb-second (V2) language like Dutch both allow clause-initial position, referred to as a ‘cognitively privileged position’ by Lambrecht (1994, p. 31-32), to be occupied by an adverbial. Both languages crucially differ, however, with respect to the information status of the clause-initial element.

According to the V2 rule, the finite verb in Dutch declarative main clauses is fixed in second position. The preverbal slot is multifunctional and may host subjects, objects or adverbials:

Table 1. V2 word order

	X	V-finite	Subject	Object	V-nonfinite	Adverbial
a.	Ze	heeft		haar debuut	gemaakt	in Parijs
b.	Haar debuut	heeft	ze		gemaakt	in Parijs
c.	In Parijs	heeft	ze	haar debuut	gemaakt	
She made her debut in Paris.						

Adverbials used for text-structuring purposes favour initial position. Depending on discourse context, placement in initial position can also serve to add emphasis, but initial adverbials may equally function as neutral, unmarked discourse links (Los, 2012).

The SVO structure of English, on the other hand, requires the subject to precede the verb. Pre-subject position may be occupied by an adverbial, but this pre-subject element tends to be more marked than its preverbal Dutch equivalent. Unmarked discourse links are typically realised by the grammatical subject of the sentence (Los, 2012). Clause-initial adverbials account for barely 15% of all adverbials in the Longman Spoken and Written English Corpus, although preferences for placement within the sentence vary according to adverbial class (Biber *et al.*, 1999, p. 763-777).

The previous chapters have shown Dutch learners to differ from native writers in the frequency with which they use adverbials clause-initially. In particular, they seem to overuse those adverbials that serve a cohesive purpose. However, the idea that this feature of advanced Dutch EFL learners' interlanguage may be the result of V2 information-structural transfer can ultimately only be confirmed or rejected by establishing to what extent it is shared by learners with other L1 backgrounds (cf. Jarvis, 2000).

This paper therefore investigates to what extent advanced Dutch learners' use of initial adverbials is shared by Francophone Belgian learners. Unlike Dutch, French has SVO word order, both in main and subordinate clauses, but use of pre-subject position is not as restricted as it is in English. Longer adverbial phrases prefer peripheral (i.e. initial or final) positions (cf. Grevisse & Goosse, 1993; Charolles, 2003; Fuchs & Fournier, 2003; Magnus, 2014). Initial position is preferred for text-structuring or light emphasis¹⁷ but may also serve a variety of stylistic purposes, i.e. to avoid a series of sentences starting with a subject, create a

¹⁷ In order to really stress an element in pre-subject position, the cleft construction 'C'est ... que' is generally used (cf. Grevisse & Goosse, 1993; Deloffre, 1967). Note that French does not have a lexical but a phrasal accent, stressing the final syllable of a group or sentence, so that cleft constructions are required for *any* element (subject, complement, adverbial) that occurs in focus.

specific rhythm, or to achieve a better balance in the sentence (in terms of the combined weight of constituents preceding and following the main verb). For example, the adverbial (in bold) in the sentence below can be placed either in initial or in final position:

(1)

- a. ***Avec l'information sur l'usage et les attitudes linguistiques,** le gouvernement peut définir la politique linguistique.¹⁸*
- b. *Le gouvernement peut définir la politique linguistique **avec l'information sur l'usage et les attitudes linguistiques.***

However, if the grammatical subject takes the form of a pronoun, 'il' in the sentence below, the adverbial is more likely to be used initially. While (2b) is not grammatically incorrect, it is slightly out of balance.

(2)

- a. ***Avec l'information sur l'usage et les attitudes linguistiques,** il peut définir la politique linguistique.*
- b. *Il peut définir la politique linguistique **avec l'information sur l'usage et les attitudes linguistiques.***

With respect to circumstantial complements, Grevisse and Goosse (1993, p. 162) note that 'it is often because of stylistic reasons (emphasis, harmony, balance

¹⁸ This roughly translates as: 'With the information about the usage and the linguistic attitudes, the government can define the linguistic policy'.

within the sentence) or the natural order of ideas that the circumstantial complement is assigned its position relative to the verb and other complements' (our translation). A similar observation is made with respect to adverbs in general: 'The position of the adverb is quite variable: it is often because of stylistic reasons – balance, rhythm, harmony, emphasis – that this word is assigned its position in the sentence' (our translation, Grevisse & Goosse, 1993, p. 870).

Unlike previous studies (cf. Granger & Tyson, 1996; Altenberg & Tapper, 1998), this paper is not limited to linking adverbials, but considers L1 Dutch, French and English novice writers' use of all types of initial adverbials. It aims to determine whether Dutch learners' tendency to place adverbials in pre-subject position is a) a transfer-induced feature of Dutch English b) a more widely-shared interlanguage feature, or c) characteristic of novice writing in general. We hope to answer this question by means of a quantitative and qualitative Contrastive Interlanguage Analysis (Granger, 1998) of argumentative texts by advanced learners of English whose L1 is either Dutch (Germanic V2) or French (Romance SVO) as well as a corpus of argumentative essays by NS novice writers.

2. Methodology

2.1 Corpora

We compared use of pre-subject adverbials in the Dutch and Francophone components of LONGDALE (Meunier, 2015), collected at Radboud University, the Netherlands, and Université Catholique de Louvain, Belgium, respectively. The A-levels in LOCNESS (Granger, 1996) were used as a reference corpus of NS novice writing. For reasons of comparability, we only considered timed argumentative essays, which meant that no year 3 material for the Dutch learners could be included. Table 2 contains an overview of texts and word counts:

Table 2. Text and word count learner and NS corpora

	LONGDALE-French		LONGDALE-Dutch		LOCNESS-NS	
	texts	words	texts	words	texts	words
Year 1	107	59920	108	28956		
Year 2	72	42600	97	26976		
Year 3	63	35823				
Total	242	138343	205	55932	114	60739

We added syntactic annotation to the corpora by means of the Stanford Parser (Klein & Manning, 2003) and subsequently filtered out all pre-subject adverbial phrases using Corpus Studio (Komen, 2012). The output, a database of 3271 pre-subject adverbials, was then manually annotated in Cesax (Komen, 2012). Each adverbial received a semantic category and referential state label, according to Biber *et al.* (1999) and Komen (2013) respectively (cf. chapter 1). For the purpose of this paper, three adverbial classes were distinguished: circumstance, stance and linking (Table 3). For each adverbial class, we paid special attention to those adverbials that have an identity link to an antecedent in the directly preceding context, referred to as ‘local anchors’ (Los, 2012), as these are likely to serve a cohesive purpose.

Table 3. Adverbial annotation scheme

adverbial class	local anchor	example
circumstance (time)	no	Not only is the number of violent films increasing, but the content is also becoming more and more violent. A dozen years ago , most directors used to avoid showing crude violent scenes. (L1 French, student UCL0015, Yr1)
circumstance (place)	yes	And this is something that is supported by the environmental policy plan mentioned before; In it , the RU seems to be well aware of the fact that the amount of waste is still relatively high (L1 Dutch, student RAD0970, Yr2)
stance (attitude)	no	Money is vital, we need it to survive. Unfortunately people often have an abusive use of money. (L1 French, student UCL0012, Yr1)
linking (contrast)	no	Some people say that high school pupils should be encouraged to take a gap year before committing to years of study at a university. However , after leaving the study environment to work or go abroad it is very difficult to start with an education again, because studying is then no longer part of one's system or everyday passtime. (L1 Dutch, student RAD0940, Yr2)

This procedure allowed us to consider not only linking adverbials included in a more or less comprehensive list based on previous research or standard grammars, but also nonstandard alternatives, including misspellings, which we would not have been able to find in the literature. It also enabled us to include circumstance and stance adverbials, which could not possibly be represented by a finite list.

2.2 Statistical analysis

The data were analysed by means of Bayesian parameter estimation (described more elaborately in chapter 4). Unlike frequentist approaches, Bayesian parameter estimation results in a posterior distribution, which relates each possible outcome to the probability of this outcome being the true value of the coefficient, given the data.

To better interpret the posterior distributions, two additional scores were computed: (1) the mean of the posterior distribution, which takes both the most likely value into account, as well as the uncertainty (or variance), by weighing each value on the x-axis by its corresponding probability density, and (2) the Probability of Equal Coefficients (PoEC), which indicates the probability that the parameter values for two groups are identical.

3. Quantitative results: frequencies of occurrence

3.1 Adverbials overall

The posterior distributions show that both learner groups start out using considerably more initial adverbials than the NSs, with (almost) completely non-overlapping distributions and 0.000 PoECs.

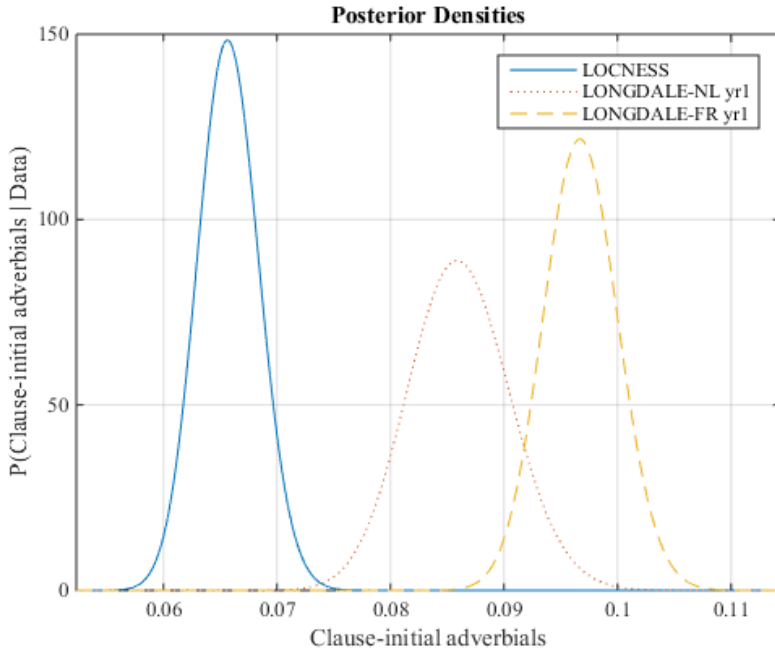


Figure 1. Posterior distributions pre-subject adverbials overall year 1

With a mean of 0.097 against 0.086 (Table 4), the Francophone learners use even more pre-subject adverbials than the Dutch. The Probabilities of Equal Coefficients (PoECs), represented in the form of a crosstable, show that in year 1 the distributions of the Dutch and Francophone learners overlap slightly, with a 0.012 PoEC (Table 5).

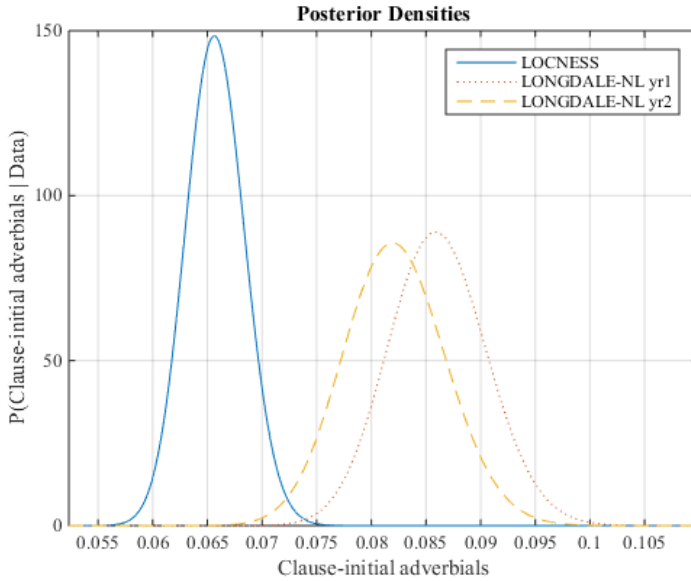
Table 4. Mean values pre-subject adverbials overall

Dutch	Year 1	0.086
	Year 2	0.082
French	Year 1	0.097
	Year 2	0.091
	Year 3	0.095
NS		0.066

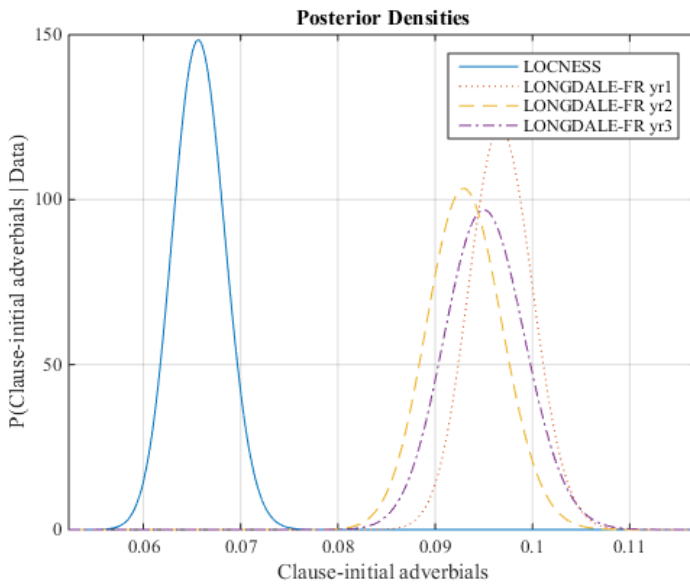
Table 5. PoECs pre-subject adverbials overall

		Dutch		French			NS
		Year 1	Year 2	Year 1	Year 2	Year 3	
Dutch	Year 1	1.000	0.051	0.012			0.000
	Year 2	0.051	1.000		0.013		0.001
French	Year 1	0.012		1.000	0.060	0.072	0.000
	Year 2		0.013	0.060	1.000	0.066	0.000
	Year 3			0.072	0.066	1.000	0.000
NS		0.000	0.001	0.000	0.000	0.000	1.000

In year 2, both learner groups are slightly closer to the NSs. However, the overlap between subsequent years within each of the learner corpora is so considerable (Figure 2) that there is no evidence of any longitudinal development. Frequencies merely fluctuate.



a. LOCNESS and LONGDALE-Dutch year 1 and 2



b. LOCNESS and LONGDALE-French year 1, 2 and 3

Figure 2. Posterior distributions pre-subject adverbials overall

Although it is noteworthy that the L2 novice writers use more pre-subject adverbials than NS novice writers, different classes of adverbials vary in their preferences for initial, medial or final position, which necessitates a separate analysis of circumstance, stance and linking adverbials.

3.2 Circumstance adverbials

Use of circumstance adverbials by both learner groups is initially largely comparable, with mean values in year 1 of 0.034 for the Dutch and 0.035 for the Francophone learners (Table 6). There is considerable overlap between both distributions, with a 0.096 PoEC (Table 7). The mean values of both Francophone and Dutch learners exceed the NS mean of 0.027 and PoECs between NS and both other groups are relatively low.

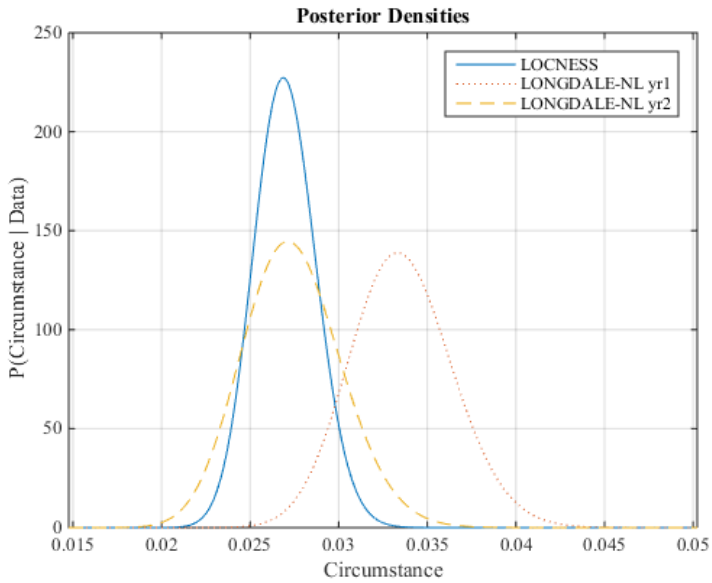
Table 6. Mean values pre-subject circumstance adverbials

Dutch	Year1	0.034
	Year 2	0.027
French	Year 1	0.035
	Year 2	0.033
	Year 3	0.038
NS		0.027

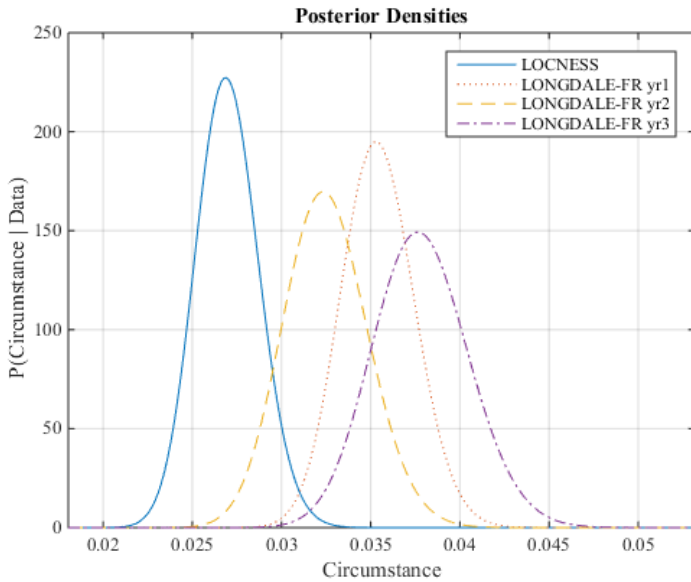
Table 7. PoECs pre-subject circumstance adverbials

		Dutch		French			NS
		Year 1	Year 2	Year 1	Year 2	Year 3	
Dutch	Year 1	1.000	0.030	0.096			0.018
	Year 2	0.030	1.000		0.041		0.121
French	Year 1	0.096		1.000	0.082	0.093	0.001
	Year 2		0.041	0.082	1.000	0.037	0.024
	Year 3			0.093	0.037	1.000	0.000
NS		0.018	0.121	0.001	0.024	0.000	1.000

There is a slight decrease in the use of circumstance adverbials by Francophone learners between year 1 and 2. Again, this does not seem to represent a steady development, considering the subsequent rise in year 3 (Figure 3). There is considerable overlap between successive years, with relatively high PoECs (Table 7). There is more development between the Dutch year 1 and year 2 mean. At 0.027, the Dutch year 2 mean is identical to that of the NSs (Table 6), but in the absence of data for year 3, it is impossible to assess whether this downward trend would have continued.



a. LOCNESS and LONGDALE-Dutch year 1 and 2



b. LOCNESS and LONGDALE-French year 1, 2 and 3

Figure 3. Posterior distributions pre-subject circumstance adverbials

3.3 Stance adverbials

For stance adverbials, a more conspicuous difference between both groups can be observed. While the Dutch initially use fewer stance adverbials in comparison with NSs, the Francophones use more (Table 8). There is no overlap between the Francophone and Dutch learners, with a 0.000 PoEC (Table 9). The NSs are situated in between both groups, not overlapping with the French and only slightly with the Dutch learners.

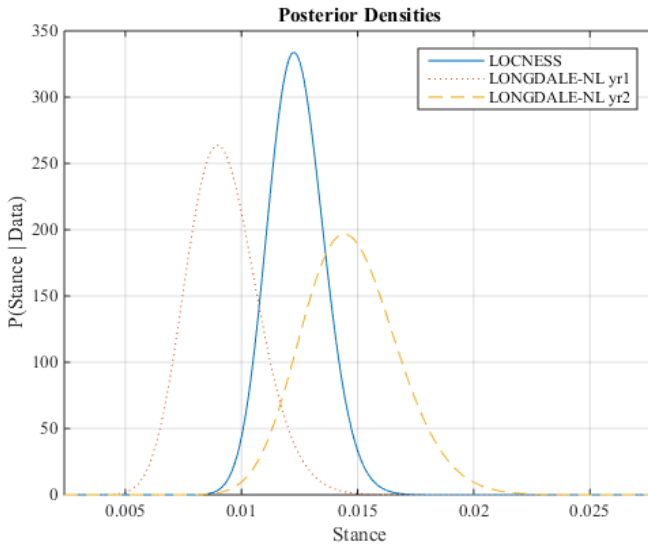
Table 8. Mean values pre-subject stance adverbials

Dutch	Year 1	0.009
	Year 2	0.015
French	Year 1	0.020
	Year 2	0.024
	Year 3	0.016
NS		0.012

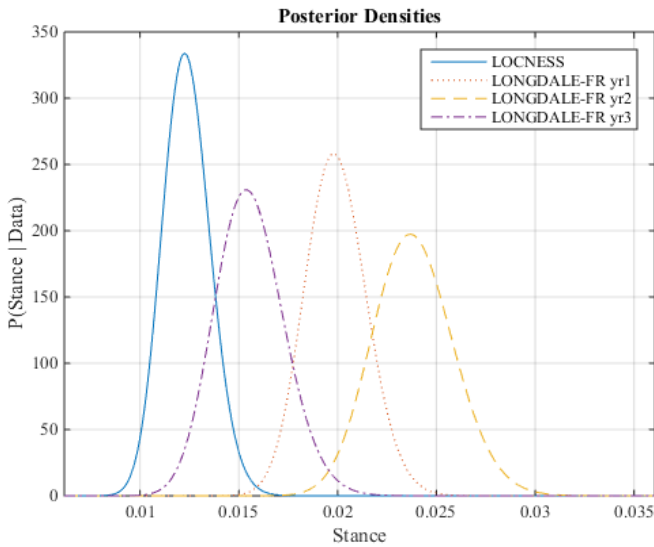
Table 9. PoECs pre-subject stance adverbials

		Dutch		French			NS
		Year 1	Year 2	Year 1	Year 2	Year 3	
Dutch	Year 1	1.000	0.016	0.000			0.056
	Year 2	0.016	1.000		0.001		0.109
French	Year 1	0.000		1.000	0.049	0.031	0.000
	Year 2		0.001	0.049	1.000	0.001	0.000
	Year 3			0.031	0.001	1.000	0.062
NS		0.056	0.109	0.000	0.000	0.062	1.000

Use of stance adverbials by both learner groups rises between year 1 and 2:



a. LOCNESS and LONGDALE Dutch year 1 and 2



b. LOCNESS and LONGDALE-French year 1, 2 and 3

Figure 4. Posterior distributions pre-subject stance adverbials

The rise in the number of stance adverbials between year 1 and 2 increases the overlap in the distributions between the Dutch learners and NSs, with a PoEC of 0.109 (Table 9), whereas the gap between the NSs and the Francophones widens. Between year 2 and 3 the Francophone learners' use of stance adverbials decreases considerably.

3.4 Linking adverbials

Both groups of learners use considerably more linking adverbials in pre-subject position than the NSs, with completely non-overlapping distributions (Figure 5) and 0.000 PoECs in both cases (Table 10). Mean frequencies in the Dutch sub-corpus slightly exceed those in the Francophone sub-corpus (Table 11), although the distributions of both groups overlap considerably, with a 0.089 PoEC.

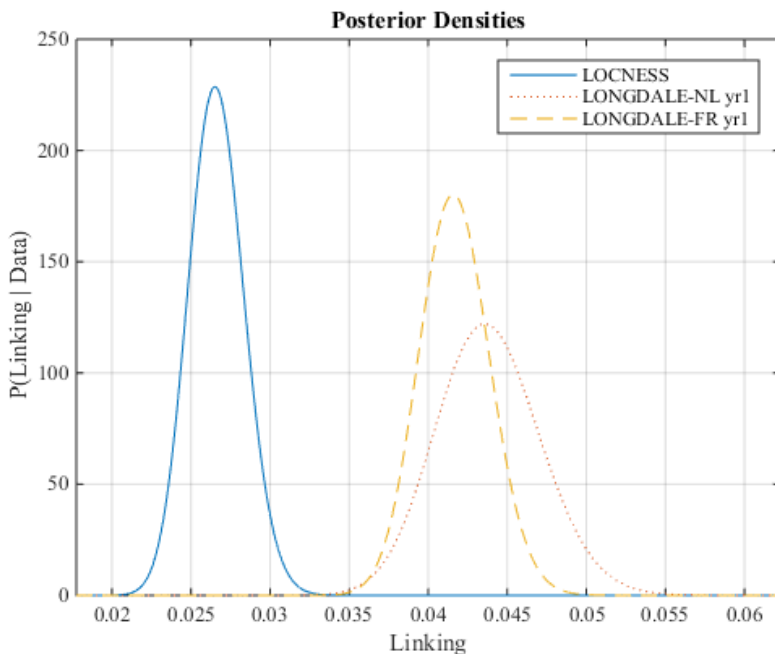


Figure 5. Posterior distributions pre-subject linking adverbials year 1

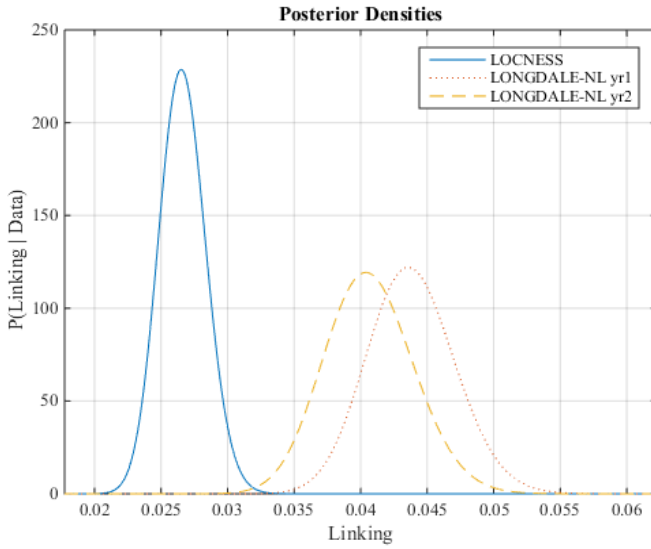
Table 10. Mean values pre-subject linking adverbials

Dutch	Year 1	0.044
	Year 2	0.041
French	Year 1	0.042
	Year 2	0.037
	Year 3	0.042
NS		0.027

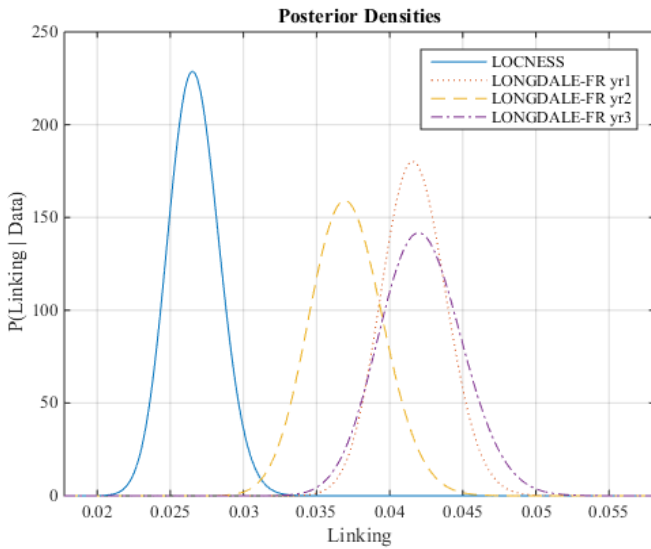
Table 11. PoECs pre-subject linking adverbials

		Dutch		French			NS
		Year 1	Year 2	Year 1	Year 2	Year 3	
Dutch	Year 1	1.000	0.068	0.089			0.000
	Year 2	0.068	1.000		0.067		0.000
French	Year 1	0.089		1.000	0.046	0.110	0.000
	Year 2		0.067	0.046	1.000	0.043	0.000
	Year 3			0.110	0.043	1.000	0.000
NS		0.000	0.000	0.000	0.000	0.000	1.000

There is a slight decrease between year 1 and 2 for both the Dutch and the Francophone learners (Table 10). By year 3, however, the Francophones have returned to their year 1 mean, suggesting that, despite the decrease between year 1 and 2, there is no consistent development. While for the Dutch learners we have no data for year 3, learner distributions in year 1 and 2 are not sufficiently distinct to assume a development is taking place.



a. LOCNESS and LONGDALE Dutch year 1 and 2



b. LOCNESS and LONGDALE-French year 1, 2 and 3

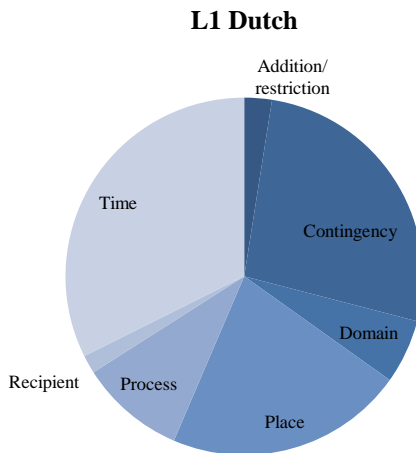
Figure 6. Posterior distributions pre-subject linking adverbials

All in all, our quantitative analysis leads us to conclude that Dutch learners of English are not exceptional in their relatively frequent use of adverbials in pre-subject position. In fact, Francophone learners appear to use even more. The three sub-corpora differ in the relative frequency of occurrence of each adverbial type. We will now zoom in on the use of these adverbials in context.

4. Qualitative results: adverbials in context

4.1 Circumstance adverbials

Figure 7 represents total use of initial circumstance adverbials by L1 Dutch, L1 French and NS novice writers, subdivided by semantic category:



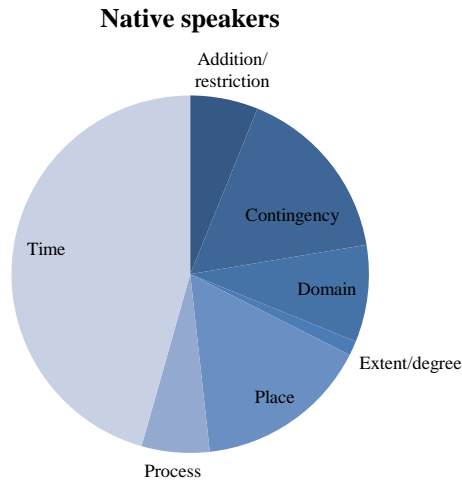
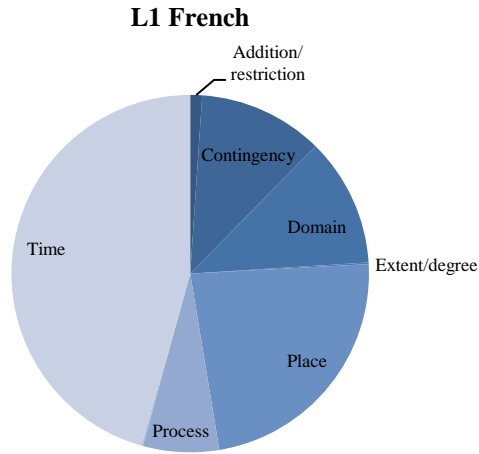


Figure 7. Distribution of initial circumstance adverbials

‘Time’, ‘Place’ and ‘Contingency’ are the largest categories of circumstance adverbials. Contingency adverbials, which the Dutch learners appear to rely on more heavily (27%) than either the NS writers (16%) or the French learners (11%),

include a wide variety of related categories that may express contrast, concession, cause or reason, condition, result, or purpose (see Appendix II, Table 1). Because of the limited size of each of these subcategories, we will here concentrate on adverbials expressing time and place.

4.1.1 *Time*

Time adverbials occur relatively more frequently in initial position than other categories of circumstance adverbials, as they have an important text-structuring function (Biber *et al.*, 1999, p. 804). It may be partly for this reason that time adverbials constitute the largest category of initial circumstance adverbials for all three groups, although the proportion of time adverbials is considerably smaller for the Dutch learners (32%) than for the French or L1 novice writers (both 46%). Initial time adverbials are typically used to mark a shift between the preceding and subsequent discourse:

- (3) It is a well-known fact that before the fifties, women were just allowed to work in their house and for the welfare of their family. But thanks to the second World War, women enter in the professional world: women constructed weapons for the soldiers. **Since that moment**, women were persuaded that they could work. **Nowadays**, women have quite established parts in the societies.

(L1 French, student UCL0042, Yr2)

Placement in initial position may also serve to give time adverbials ‘added salience’ (Biber *et al.*, 1999, p. 804). The placement of ‘already’ in the following example coincides with its emphatic use:

- (4) If they don't build more roads then people complain about the density of traffic and if they do then rabid environmentalists spring out from their underground and start chaining themselves to trees! Just look at the case of the planned Newbury bypass. **Already** it is behind schedule due to the ranting of these 'environmentalists'. Even the most ardent of these would acknowledge the necessity of something being done on this stretch of road.

(L1 English, A-levels1_transport06)

It appears that not all Dutch and French learners are aware of the added salience that placement in initial position may entail. In the following example, placement in initial position gives the adverbial a prominence that does not fit the essentially neutral time indication it conveys:

- (5) But of course, to become this rapid beauty, self-confidence is presumed. This not being the normal. **At some points in their life**, everybody in this world has doubts. Surely we could not assume that because they are not certain of themselves, they cannot order some people.

(L1 French, student UCL0088, Yr2)

Similarly, the positioning of the adverbial in the following sentence unintentionally gives it added salience, perhaps even a hint of contrastiveness, especially in comparison with the more neutral option of realising the continued topic that 'an internship' represents as the grammatical subject of the sentence (e.g. 'an internship allows the student to ...').

- (6) When the applicant also has a great reference from his or her internship it will be far more likely for that person to be hired. Secondly, **during an internship** a student can start networking.

(L1 Dutch, student RAD0004, Yr2)

This use of an initial adverbial to ‘anchor’ a topic to an antecedent in the preceding discourse, as in (6), is a strategy that is used relatively more frequently by the Dutch learners than by either other group. Unlike many other time adverbials, local anchors are not primarily used for overall text-structuring purposes, but serve to create cohesion locally. They are realised as prepositional phrases, consisting of a preposition followed by a pronoun or noun phrase that refers to an antecedent in the directly preceding sentence:

- (7) Moreover, students also get the opportunity to prepare for an exam to be able to score high grades, but this does not necessarily mean that the information will linger longer in their minds. **During an exam** they will write down everything they can remember, but once they have written it down, the information will instantly disappear from their memories.

(L1 Dutch, student RAD1088, Yr2)

Of all time adverbials realised as prepositional phrases, Dutch learners use 11% for local anchoring, compared with 4% for the NSs and 5% for the Francophones.

A final difference in the use of time adverbials by the different groups is that Francophone learners prefer to realise time adverbials as single adverbs such as ‘sometimes’, ‘now’ and ‘nowadays’. The use of adverb phrases is not as frequent but includes phrases such as ‘years ago’ and ‘a long, long time ago’. Prepositional phrases, which tend to be longer, constitute the most common realisation of time adverbials for both other groups of novice writers:

Table 12. Realisation of initial time adverbials

	prepositional phrases	single adverbs and adverb phrases	noun phrases
NS	54%	36%	10%
Francophone learners	35%	59%	6%
Dutch learners	63%	27%	10%

Biber *et al.* (1999, p. 808) note that single adverbs occur much more frequently in spoken discourse than in fiction, news or academic prose. The realisation of time adverbials by Francophone learners might therefore point towards a lack of register awareness.

4.1.2 *Place*

English place adverbials, mostly realised as prepositional phrases, have a strong preference for final position (Biber *et al.*, 1999, p. 787, p. 802). The proportion of initial place adverbials is bigger for both learner groups (23% and 22% for Francophone and Dutch learners respectively) than for the NS novice writers (16%). One of the functions that initial place adverbials fulfil is that of general scene setting, often commenting on the state the world is in:

(8)

- a. **In a world where everybody is supposed to be concerned about the planet and global warming**, this use of disposables is somewhat odd.

(L1 Dutch, student RAD0010, Yr2)

- b. **In a world where books are increasingly replaced by movies** it would be quite surprising if nobody cared about how movies put limits to our imagination.

(L1 French, student UCL0123, Yr3)

Place adverbials are also used to announce the topic of the essay:

(9)

- a. **In this essay**, I wish to show that in our modern world, dominated by science, technology and industrialization, there is no more place for dreaming and imagination.

(L1 French, student UCL0008, Yr1)

- b. **In this assignment**, I wish to show that lying can be both positive and negative.

(L1 French, student UCL0006, Yr1)

Adverbials of the kind exemplified in (9) are used relatively often (21% of all place adverbials) by the Francophone writers, but only incidentally (8%) by the Dutch, and never by the NSs. This may partly reflect a difference in writing culture, although teaching is likely to play a role as well. These specific groups of Dutch learners were explicitly and repeatedly told in class to include a thesis statement in their argumentative essays rather than a statement of intent.

Like time adverbials, place adverbials are often used initially in L1 English to highlight or contrast a specific element. Consider for instance the examples in (10), where contrast is either explicitly stated or implied:

(10)

- a. Since the beginning of man on this earth we have always detected, stalked, captured and killed our prey. This was essential for our survival.

In the modern world there is no need to hunt in order to obtain our food.

(L1 English, A-levels3_fox_hunting04)

- b. Hopefully this would lead to a reversal of the damage we've done to the environment through reduced vehicle emissions. We should also encourage cycling as a mode of transport. **In Holland** there is a complete network of cycle paths as well as roads.

(L1 English, A-levels1_transport03)

Similar place adverbials in the English texts produced by L1 Dutch writers may seem to have been used to achieve the same effect, but a contrastive reading is ruled out by the context:

(11)

- a. The Dutch government instead wants to spend more money on primary education because this is much more important than university. Primary education is the basis for further education and it is obligatory for all children living in the Netherlands. **In the Netherlands** you cannot leave school legally till you are eighteen.

(L1 Dutch, student RAD1111, Yr1)

- b. The goal of primary education is achieving basic skills such as learning how to read and write for all the pupils. Other subjects that the young children are taught in primary education, such as mathematics, geography, history are about as important. Also, **in primary education**, the children develop creative skills in subjects such as music, drawing and physical education. All of these subjects cover a very broad field, and it lays out the basis for an entire generation of young kids that they will need for the rest of their lives. **In primary school**, they will also develop most of their social skills, as they will be forced for the first time in their lives to work together with other pupils of their age.

(L1 Dutch, student RAD1129, Yr1)

The adverbials in (11) continue the topic of the preceding discourse. It appears, therefore, that they are used as neutral rather than contrastive discourse links. These adverbials are syntactically correct in both Dutch and in English, but the pragmatics of this construction in the Dutch learners' L1 and L2 do not coincide. They may therefore unintentionally receive a marked reading.

In L1 English writing, initial place adverbials may also serve a cohesive purpose by 'using some information given in the previous discourse as the starting point for the next sentence' (Biber *et al.*, 1999, p. 809). In Francophone English writing, place adverbials can also be used for local anchoring, but there is a demonstrable difference in frequency between the three language groups. Whereas 25% of the initial place adverbials used by the Dutch learners are used for local anchoring, this applies to only 9% and 3% of the place adverbials used by NS and Francophone novice writers respectively.

4.1.3 Adverbial length

Apart from semantic category, weight is also likely to affect adverbial placement in the sense that in English long adverbial phrases may be expected to favour final rather than initial or medial position. This applies to circumstance adverbials in particular, as they are mostly realised as prepositional phrases and allow for greater variability in length than most stance and linking adverbials (Biber *et al.*, 1999, pp. 768-770). Transfer might be expected to lead both Dutch and French learners to produce longer initial circumstance adverbials than NSs. However, this expectation is not corroborated by an analysis of the mean length of the adverbial categories discussed in this section, i.e. time and place (Table 13). Even if prepositional phrases (PPs) are considered separately from adverb phrases (APs) and noun phrases (NPs), mean values for each of the three language groups are very close. Only the temporal noun phrases produced by the NSs are longer than those of L1 Dutch and L1 French writers due to the frequent inclusion of 'each/every time' followed by a relative clause.

Table 13. Mean length of initial time and place adverbials (in number of words)

		L1 Dutch	L1 French	NS
Place	PP	4.02	4.14	3.97
	AP	2	1.1	1
Time	PP	4.36	4.26	4.64
	AP	1.25	1.34	1.63
	NP	2	2.77	4

Clearly, the three groups can hardly be distinguished on the basis of mean adverbial length. Still, there is a demonstrable difference between L1 French and L1 Dutch learners on the one hand and NS writers on the other in the occurrence

of 'peak values'. These peaks take the form of prepositional phrases consisting of as many as 20 words, far exceeding anything found in the NS data:

(12)

- a. **With the increasing number of mass medias and the invention of television and internet and the decreasing price of information**, some people, like actors, musicians and sportsmen, were known by a large number of people.

(L1 French, student UCL0007, Yr2)

- b. **In spite of the fact that all those progresses have mainly good effects and are, in general, very usefull**, they have also their part of disadvantages and nefast effects.

(L1 French, student UCL0017, Yr3)

- c. **Because of the wide arrange of choices the students have, and the fact that they should be supported equally regardless of the choice they make**, this requires quite an extensive education network.

(L1 Dutch, student RAD1185, Yr1)

In the French data, similarly top-heavy sentences result from the placement of strings of adverbials at the beginning of a clause, punctuated by commas:

(13)

- a. **Furthermore, in 1986 approximately, in Russia** there was another explosion which killed thousands of people.

(L1 French, student UCL0040, Yr1)

- b. **Nowadays, in this terribly active, or even competitive world we live in,** a job is something every single person has to have.

(L1 French, student UCL0130, Yr2)

- c. **But nowadays, with all the new ways of communication, especially thanks to internet,** you can be aware of what it is happening everywhere and you pay more attention to unworthy thing like to know if Britney Spears got a new haircut or not.

(L1 French, student UCL0026, Yr2)

In many cases Francophone learners' use of strings of initial circumstance adverbials appears to mimic the kind of rhythmic structure commonly found in written French.¹⁹ Similar strings do not occur at all in the Dutch EFL data. This is perhaps not surprising given the fact that Dutch is a V2 language which does not formally allow more than one constituent to be inserted in the preverbal slot.

¹⁹ This rhythmic structure is exemplified in sentences such as the following, in which a string of initial adverbials, punctuated by commas, precedes the subject: '*De façon récurrente, par compassion, par chagrin, en écoutant la chanteuse Aretha Franklin ou pour remercier son équipe de campagne, le président américain n'hésite pas à laisser émerger son émotion*'. This roughly translates as 'Repeatedly, because of compassion, because of sorrow, when listening to the singer Aretha Franklin or thanking his campaign team, the American president does not hesitate to show his emotion' ('Larmes politiques'. *Le Monde* 16-1-2016).

4.2 Stance adverbials

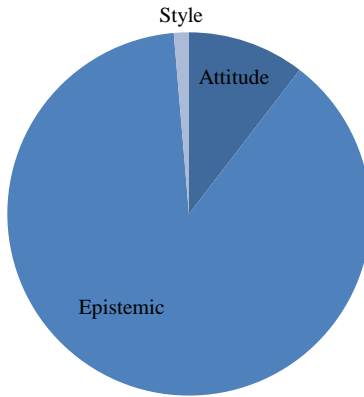
Overall, stance adverbials are considerably less frequent than circumstance adverbials and have a strong preference for medial position (Biber *et al.*, 1999, p. 853, p. 872). Certain stance adverbials do not only comment on the content or style of a proposition, which is their primary function, but also serve a cohesive purpose, in which case they occur more often in initial position (p. 874). This is true for adverbials such as ‘in fact’ or ‘indeed’, which reinforce the preceding context. Biber *et al.* (1999) divide stance adverbials into three main semantic categories: (1) epistemic adverbials, which ‘express the speaker’s judgment about the certainty, reliability, and limitations of the proposition’, such as ‘undoubtedly’ or ‘in fact’ (p. 854), (2) attitude adverbials, which ‘typically [convey] an evaluation, value judgment, or assessment of expectations’, such as ‘fortunately’ or ‘hopefully’ (p. 856), and (3) style adverbials, which ‘comment on the manner of conveying the message’, such as ‘honestly’ or ‘frankly’ (p. 857). The proportions of the different categories per group are given in Figure 8.²⁰

²⁰ Stance adverbials are not generally used for local anchoring. The Dutch and Francophone learners use 4.55% and 3.21% respectively to refer back to an antecedent in the directly preceding discourse. In the NS novice writing corpus, referential use of stance adverbials does not occur at all.

L1 Dutch



L1 French



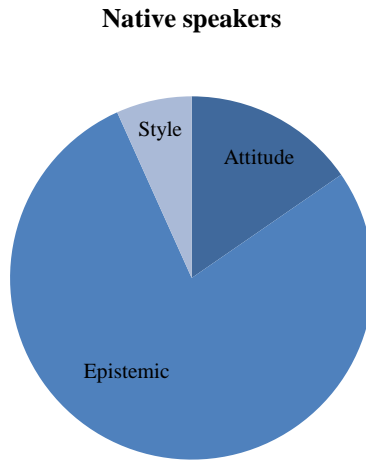


Figure 8. Distribution of stance adverbials

Stance adverbials are typically fixed expressions, often in the form of single adverbs or short PPs, which do not allow for a lot of variation. The analysis in the previous section revealed that French learners use more stance adverbials overall than either other group. This is mainly due to French learners' frequent use of epistemic stance adverbials, which can be further subcategorised into 'actuality and reality', 'doubt and certainty', 'imprecision', 'limitation', 'viewpoint or perspective' and 'source of knowledge' (cf. Appendix II, Table 2). The most notable differences in distribution and realisation between the different language groups occur in the use of epistemic stance adverbials expressing 'viewpoint or perspective' and 'actuality and reality'. We will discuss these in the following subsections.

4.2.1 *Viewpoint or perspective*

Viewpoint or perspective adverbials are used more often by both groups of L2 writers than the L1 English writers. This category, including adverbials such as 'in

my opinion' and 'in my view', accounts for 15% of all stance adverbials by the NS novice writers, against 26% and 22% for the Francophone and Dutch EFL writers respectively. The Francophone writers do not only differ from the NSs in frequency, but also in the range of tokens used in this category. The NSs essentially use these adverbials to introduce their own viewpoint (e.g. 'in my opinion', 'in my eyes'), whereas the Francophone learners use them more productively, and typically also with a wide variety of referential entities:

(14)

- a. Some years ago and it is not so far actually it was well define in the people's mind that men went to work and women stayed home to take care of the children and also to take care of the house. **In that family life point of view**, the men were not often at home and they didn't often see the children, they didn't see them growing up.

(L1 French, student UCL0096, Yr2)

- b. I personally think that it depends on what we consider. **On a scientific and technological point of view** it is clear that people could not dream of anything better that what we possess in our actual world.

(L1 French, student UCL0075, Yr1)

It is also interesting to note the frequent use of 'according to [+NP]' in the L1 French corpus. While this PP is generally used to introduce a source of knowledge, 47% of all occurrences in the Francophone corpus are used to mark the author's own point of view:

- (15) This proves that still there are progress to do on this subject. **According to me**, progress firstly need to come from men 's opinion about women.

(L1 French, student UCL0062, Yr2)

‘According to’ occurs only twice in the L1 Dutch corpus, in both cases as a neutral introduction of a source of knowledge, and only once in the L1 English corpus, where it is used contrastively, to distance the writer from the message in the remainder of the clause:

- (16) But when it comes to ceasing the hunting and brutal killing of one of Britains few natural animals they seem to turn a blind eye. **According to people who partake in fox hunting** it is an enjoyable sport.

(L1 English, A-levels3_fox_hunting03)

While one of the occurrences of ‘according to [+NP]’ in the L1 French EFL corpus implies a similarly critical view, most are essentially neutral:

- (17)
- a. Violence is not especialy, as many people can believe, an adult problem like war or terrorism. **According to current experiences**, children are getting more and more agresive.

(L1 French, student UCL0112, Yr1)

- b. **According to a general view**, we could say self-confidence is a major requirement for reaching one’s goals.

(L1 French, student UCL0134, Yr2)

Compared with most other stance adverbials, 'according to [+NP]' occurs relatively more frequently in initial position, at least in written registers (Biber *et al.*, 1999, p. 873). What distinguishes the L1 French EFL novice writers from both other groups, therefore, is not likely to be the result of different preferences for placement. Instead, it is the variety of functions that these adverbials serve in Francophone EFL writing that sets them apart.

4.2.2 *Actuality and reality*

Frequencies of actuality and reality adverbials (e.g. 'in fact', 'indeed') also vary greatly. In comparison with the NSs, whose proportional use of actuality and reality adverbials is 13%, the L1 Dutch writers use very few. Only 5% of their stance adverbials belong to this category. The L1 French writers, on the other hand, systematically punctuate their texts with actuality and reality adverbials. At 27%, they use considerably more compared with either L1 Dutch EFL writers or L1 English writers.

In line with observations by Granger & Tyson (1996, pp. 20-22), our L1 French data show that the adverbials 'in fact' and 'as a matter of fact' are often deployed as equivalents of French '*en fait*', which serves to reformulate or synthesise preceding information. In English, 'in fact' and 'as a matter of fact' also have a secondary cohesive purpose, as they 'mark not just the nature of the clause, but also its connection to the previous discourse' (Biber *et al.*, 1999, p. 874). However, rather than introduce a reformulation only, they also add something new or surprising to the preceding information. The following examples clearly illustrate this difference:

(18)

- a. Biological warfare is not the scientists fault. He did not use it to kill. **In fact** it was probably an experiment with herbicides which went wrong.

(L1 English, A-levels8_icle-alev-0027)

- b. But with all the industrialisation, everybody is seen as having the same ideas, the same point of view. Everybody buys the same products, wears the same clothes, eats the same food, drives the same car. **In fact**, everybody has the same behavior.

(L1 French, UCL0084, Yr1)

Similar transfer effects can be observed for L1 French EFL writers' use of 'actually' as an equivalent for '*actuellement*' as well as 'indeed' for '*effectivement*' or '*en effet*'. Like 'in fact', 'indeed' is often used by the L1 English writers to introduce extra detail to support the proposition in the preceding sentence, while in the Francophone EFL corpus it is predominantly used for reformulation:

(19)

- a. These factors could have a considerable impact on the agricultural industry preventing the sale of British dairy livestock not only in Britain but for export as well. With global communications as good as they are now, it could send a scare across the world about British Beef - **indeed** France and Germany already have restricted and some places banned the sale or import of British beef for fear of it's safety.

(L1 English, A-levels9_icle-alev-0007)

- b. Another benefic point in the fact that mothers stay at home, is that it frees jobs for other people. **Indeed**, if more women stay home, there will be more jobs available and so less unemployment.

(L1 French, student UCL0034, Yr2)

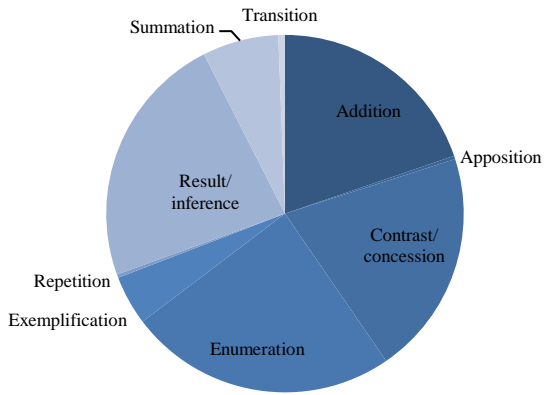
As noted by Granger and Tyson (1996, pp. 22-23), this use is very likely reinforced by lists of linking devices, readily available in textbooks and on the internet, which present these French and English terms as equivalents, without further specifying their use.

4.3 Linking adverbials

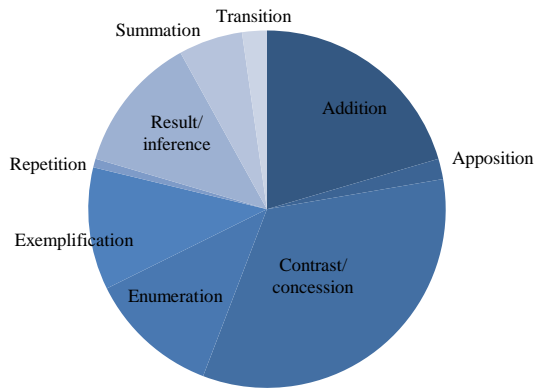
Linking adverbials function as signposts of logical relations such as addition, contrast, concession and exemplification, and are typically expressed by single adverbs or short, mostly fixed, prepositional phrases. Contrary to circumstance and stance adverbials, 'initial position can ... be considered the unmarked position for linking adverbials' (Biber *et al.*, 1999, p. 891), as they serve to mark a connection with the preceding discourse. Raw frequencies are given in Appendix II (Table 3). Figure 9 presents an overview of the proportional use of the different categories of linking adverbials.²¹

²¹ Like the stance adverbials discussed in the previous section, the linking adverbials are predominantly used non-referentially. The Dutch use them for local anchoring in only 3.29% of all occurrences of initial linking adverbials, the French in only 0.92%, and the NS novice writers in LOCNESS never use them for such a purpose at all.

L1 Dutch



L1 French



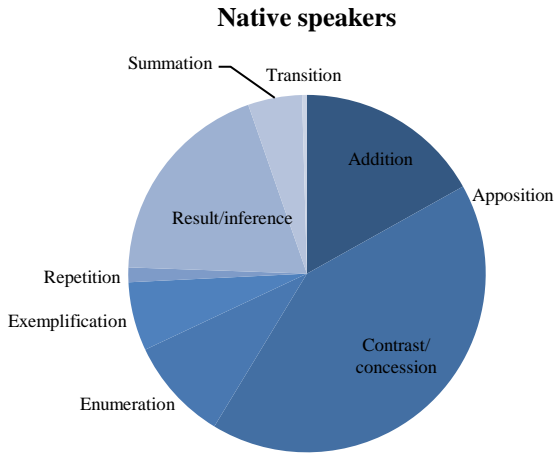


Figure 9. Distribution of linking adverbials

We will now zoom in on the categories of addition and contrast/concession, which together account for the majority of linking adverbials used by NS and Francophone novice writers.²²

4.3.1 *Addition*

Differences between L1 Dutch, French and English writers in the proportional use of addition adverbials are relatively small, at 17%, 20% and 20% respectively. However, differences can be observed in their realisation. Use of ‘also’, followed by a comma, predominates in the L1 English texts, accounting for no less than 72% of all addition adverbials, against 30% for the L1 Dutch and only 3% for the L1 French writers. Francophone writers favour the use of ‘moreover’ (44%), more so

²² Enumeration adverbials and result/inference adverbials used by the three groups differ from each other in their proportional frequency of use, but not in terms of the types of tokens that are used.

than the Dutch (11%) and, particularly, the NSs (3%). Table 14 presents the frequency of the top four addition adverbials:

Table 14. Top four addition adverbials

	L1 English	L1 French	L1 Dutch
Also	71%	3%	30%
Moreover	3%	44%	11%
Furthermore	3%	25%	23%
In addition	3%	11%	12%
Other	20%	17%	34%

The frequent use of ‘moreover’, ‘furthermore’ and ‘in addition’ by the L1 French writers and, to a lesser extent, by the L1 Dutch writers may be a teaching effect, as these adverbials invariably feature in textbooks, in lists of additive adverbials as more formal equivalents of other connectors. Their frequency may be explained by their perceived enhancement of register-appropriacy. While in L1 English ‘moreover’ and ‘furthermore’ are typically used to express an additive logical relationship between two points, in which the second is stronger or more significant (Rundell, 2007, p. 4), Francophone learners tend to use them as equivalents of neutral connectors, and alternate their use:

- (20) Also with technology, video games are more and more designed and children and teenagers choose rather that than books which nevertheless call for imagination. **Furthermore** I don't think that staying for hours in front of a screen is really better for the health whereas playing outside or reading a book is not unhealthy at all. **Moreover** video games can be violent and that could have consequences on the players.

(L1 French, student UCL0026, Yr1)

L1 Dutch use of addition adverbials stands out for different reasons. While most other linking adverbials are essentially non-referential, addition adverbials are often used by Dutch learners to explicitly refer back to an antecedent in the preceding context:

(21)

- a. By describing critical thoughts such as these the audience will notice that your writing is objective. **On top of that**, your reliability might improve when you doubt your own opinions or conclusions.

(L1 Dutch, student RAD1059, Yr1)

- b. Higher education, on the other hand, is very expensive. Students have to pay a lot just to be able to attend classes. And **besides that**, they also have to pay for a room, food, clothes and other important necessities.

(L1 Dutch, student RAD1117, Yr1)

The prepositional phrases that are used for this purpose, mostly 'on top of that', 'next to that', and 'besides that', point to transfer from Dutch pronominal adverbs such as '*bovendien*', '*daarbij*' and '*daarnaast*'. Of all addition adverbials, 17% are used referentially by the L1 Dutch writers, against 3% by the L1 French and 0% by the L1 English.

4.3.2 Contrast/Concession

This category includes common adverbs such as 'however' and 'nevertheless' as well as fixed prepositional phrases such as 'on the contrary', and 'on the other hand'. Differences in the use of these adverbials between the three groups lie in their proportional frequency as well as lexical transfer which appears to manifest itself in the realisation of contrast/concession adverbials in the learner corpora.

For the NSs, contrast and concession adverbials constitute the single most dominant category of linking adverbials, with a proportional use of 42%. At 20%, Dutch learners' use of adverbials belonging to this category is much lower. Francophone learners' use is situated in the middle, with a proportional use of 34%.

L1 influence can be observed in the use of 'on one side' and 'on the other side', which are occasionally used by the Francophone learners, and, sporadically, by the Dutch.

(22)

- a. Last, money is also a big issue in this debate about the importance of primary and tertiary school. **On one side**, if university gets even more expensive, there is a high probability that even more people will quit attending, or not even start attending university at all.

(L1 Dutch, student RAD1133, Yr1)

- b. Violent films should be banned from TV in order to protect more the young viewers. **On the other side**, we are not sure that to ban violence in films will protect people. The truth is that violence is everywhere, violence has now taken a place in our lives.

(L1 French, student UCL0074, Yr1)

Granger and Tyson (1996, pp. 22-23) note that Francophone learners equate the adverbial 'on the contrary' with both French '*au contraire*' and '*par contre*'. Francophone learners therefore use it as an alternative for both 'however' and 'on the other hand', as exemplified in (23). Our LONGDALE data show that this use is not restricted to Francophone learners, but shows up in L1 Dutch learner writing as well (24).

(23) In my opinion, if a husband works and earns enough money for his family to live comfortably, his wife should grip the opportunity of spending more time with her children. I really believe that a motherly, as well as fatherly, presence is essential for children because they will develop confidence in their parents. **On the contrary**, if parents are absent, children will feel estrangement for them and will not be apt to communicate their joys and fears with them.

(L1 French, student UCL0039, Yr2)

(24) This gives teachers and supporting personnel the opportunity to spend more time with the children and teach them the proper skills. This will prepare them better for secondary education, and they will be more successful in acquiring other skills that they are going to need later in life. **On the contrary** the importance of secondary education should not be overlooked.

(L1 Dutch, student RAD1143, Yr1)

4.3.3 *Stylistic coherence*

More generally, L1 French writers differ from both other groups in their mixed use of formal linking elements such as ‘nonetheless’ and ‘consequently’ and informal linking adverbials such as ‘anyway’ and ‘by the way’, as in (25):

- (25) Because of our massconsumption habits we are always asking for more. That is why all kind of films, series or documentaries can be showned on television. **Nonetheless** research has showned that violent films can have a negative impact on people, especially on children and teenagers. [...] Teenagers are not always aware of the fact that what they see in films is not the reality. They can firmly believe that violence is a normal way of solving problems. **By the way** it is not the case of all teenagers but a tiny part of them, who suffered mainly from social contacts or violent behavior, should be put at risk.

(L1 French, student UCL0128, Yr3)

This confirms the observation by Granger and Tyson that ‘learners seem unaware of stylistic restrictions’, which may be the result of the way lists of apparently interchangeable linking devices are presented in textbooks (1996, pp. 23-25). On the basis of the learner data used for the present study, it seems that limited register awareness is more prevalent among the Francophone learners than the Dutch.

5. Discussion/conclusion

The preceding sections have described the use of adverbials in pre-subject position in English novice writing, and compared writers whose L1 is Dutch, French and English respectively. A comparison of these specific language groups allowed us to examine to what extent frequent use of initial adverbials might be a) a V2 transfer effect that can be observed in the interlanguage of Dutch EFL learners but not in the interlanguage of EFL learners whose L1 is a romance SVO language, b) a more widely shared interlanguage feature that manifests itself in EFL writing irrespective of the author’s L1, or c) a feature of novice writing in general, shared by L1 and L2 writers alike. Our results demonstrate that it is not only Dutch learners who start many of their sentences with an adverbial. In fact, Francophone

learners have an even higher mean use of initial adverbials. L1 English novice writers, on the other hand, use considerably fewer initial adverbials.

This might be taken to suggest that frequent use of initial adverbials is not a defining characteristic of novice writing in general or a V2 transfer effect but rather an interlanguage feature shared by Francophone and Dutch learners alike. Nevertheless, a closer analysis of circumstance, stance and linking adverbials reveals clear preferential distributional patterns for each learner group: in comparison with their Dutch peers, the L1 French learners use more circumstance and stance adverbials, whereas the use of linking adverbials is more equally distributed. A qualitative analysis also points towards the effects and potential interplay of L1 transfer, writing culture and teaching practice.

In English, French and Dutch alike, an adverbial may precede the subject of a clause, but the additional (syntactic, pragmatic and stylistic) principles that govern the specific use of such constructions in the three languages differ. In English, placement of adverbials varies with adverbial class, semantic category and adverbial length. Circumstance adverbials, place adverbials in particular, are less likely to be used initially than stance or linking adverbials. If they are, their initial position often serves to give them added salience. In Dutch, adverbials are frequently used in initial position for text structuring and local anchoring, and can be either marked or unmarked. In French, initial adverbials are commonly used for text-structuring, but they may also be placed initially for stylistic reasons, i.e. to avoid a series of sentences starting with a subject, to achieve the rhythmic structure typical of written French or to achieve a better 'balance' by redistributing the weight of the sentence, particularly if it has a relatively light subject. For both learner groups, the LONGDALE data revealed that their L1 habits are transferred into their L2 English.

Lexical transfer can be observed in L1 French learners' use of stance adverbials such as 'in fact' as a full equivalent for French '*en fait*'. This type of transfer is not as frequent for the L1 Dutch learners. More subtle traces of L1 transfer occur at the syntax-pragmatics interface: Dutch learners have a stronger tendency than either other group to use initial adverbials for local anchoring, i.e. to achieve cohesion locally by referring back to an antecedent in the directly preceding context. This mirrors a common cohesive strategy in L1 Dutch texts. Francophone learners' use of initial adverbials, on the other hand, appears to be influenced by considerations of sentence rhythm and balance. The fact that in Dutch and in French initial adverbials may function as neutral discourse links or neutral stylistic 'counterweights' respectively at times results in unintentional added salience of initial circumstance adverbials in both Dutch and Francophone learners' L2 English.

Teaching rather than transfer may lie at the root of both Francophone and Dutch learners' use of linking adverbials, both in terms of frequency and realisation. Our data offer further support for the suggestion that frequent use of initial linking adverbials may be a general teaching-induced interlanguage feature. Teaching practice may also very well offer an explanation for L1 French learners' habitual use of addition adverbials such as 'moreover' and 'furthermore', both of which are particularly low-frequent items in NS novice writing of the same genre and register, which tends to prefer the use of 'also'. Francophone learners' mixed use of formal and informal linking adverbials similarly hints at a lack of register awareness. It is tempting to attribute this to the practice of presenting learners with lists of near-equivalent items without context, as observed in chapter 5 in the context of Dutch pre-university education, but whether or not L1 French learners' use of linking adverbials is equally determined by teaching practice and textbooks would need to be confirmed by further research.

To a certain extent, L1-coloured language use is likely to disappear with increasing proficiency. This is particularly true for the lexical transfer that appears to affect the use of stance adverbials by the L1 French learners in our corpus. In this context, it would be interesting to see to what extent the occurrence of the different types of transfer observed, i.e. lexical and/or pragmatic, correlates with an objective measure of learners' proficiency. This is something that we have not been able to take into account as comparability of both learner groups was based on institutional status only. Also, a comparison with a corpus of expert L2 writing would serve to clarify to what extent the hypothesised teaching and transfer effects are reduced with increased proficiency and/or writing experience. Within the relatively limited time-frame of this study, L2 writers' comparatively high use of initial linking adverbials remains remarkably consistent.

While this study leads to new, and interesting, questions, it also takes us one step closer towards understanding the process underlying the acquisition of writing skills in a foreign language. EFL writing by different L1 groups may ostensibly be the same in writers' heavy reliance on adverbials in pre-subject position, but at the same time markedly different in the way these adverbials are realised and the pragmatic purposes they serve.

7. Conclusion

7.1 Summary

This thesis offered a contrastive and developmental analysis of Dutch learners' use of clause-initial adverbials so as to determine to what extent it may be considered a defining feature of advanced Dutch learners' written English at different stages in their interlanguage development. It started out by answering the following research questions:

1. In what way do advanced learners differ from NSs in the frequency and realisation of clause-initial adverbials?
2. How do a) proficiency level and b) longitudinal development affect Dutch learners' use of initial adverbials?

Apart from offering a descriptive account, this thesis also aimed to look into the possible underlying causes of Dutch learners' idiosyncratic use of initial adverbials, as reflected in research questions 3 and 4:

3. Is Dutch learners' use of initial adverbials likely to be transfer-induced?
 - a. Is there intra-L1-group homogeneity in their interlanguage performance?
 - b. Is there inter-L1-group heterogeneity between L1 Dutch and L1 French learners?
 - c. Is there intra-L1-group congruity between their L1 and interlanguage performance?

4. Is Dutch learners' use of initial adverbials likely to be teaching-induced?

The first two questions were addressed in chapters 2, 3 and 4, while chapters 5 and 6 considered the role of transfer and teaching in Dutch learners' characteristic use of initial adverbials. A detailed account of the findings presented in each of these chapters follows below.

7.1.1

Chapter 2 reported on a case study investigating the use of initial adverbials in advanced Dutch learners' timed in-class writing assignments and untimed literature essays as well as NS published research articles in the field of linguistics and literature. Differences between Dutch and English in the information status of clause-initial position were hypothesised to lead advanced Dutch learners to use comparatively more initial place and addition adverbials, both of which are commonly used in Dutch to provide a link to the preceding discourse. Mean values per 1,000 words turned out to vary not only per group (NS vs NNS) but also per text type. Literature essays included many more place adverbials, while addition adverbials were used considerably more frequently in non-literary in-class assignments. Similarly, the NS linguistics articles included both more place and addition adverbials than the NS literature articles and generally appeared to be more overtly structured. As expected, the NNSs differed from the NSs in that their literature essays included both more place and addition adverbials than the NS literature articles. They also used more of their place and addition adverbials for local anchoring.

Place and addition adverbials turned out to have distinct patterns of longitudinal development in the Dutch learner corpus: while there was a steady decrease in the direction of NS writing in the use of initial place adverbials

between year 1 and 3, the number of addition adverbials actually moved in the opposite direction, up and away from the NS mean. This was hypothesised to be a reflection of the fact that place and addition adverbials have different preferences for placement in L1 English as well. English place adverbials have a strong preference for final position, something which extended exposure between year 1 and 3 seems to make learners increasingly aware of. Addition adverbials, on the other hand, naturally prefer initial position as they are used for text-structuring purposes, which explains the absence of a downward trend in the use of initial addition adverbials in the learner texts. The increase in their use may be linked to learners' growing awareness of the importance of text-structuring, coupled with an over-reliance on the use of one familiar cohesive strategy. For addition adverbials, a development in the direction of native writing could only be observed if they were subcategorised according to referential state, which brought to light a steady decline in the number of addition adverbials that were used for local anchoring. Even at the end of year 3, a gap remained, as the addition adverbials in the NS corpus were hardly ever used to explicitly refer back to the preceding context.

7.1.2

Rather than taking a longitudinal perspective, the study reported on in chapter 3 analysed the correlation between proficiency and use of initial adverbials in a cross-sectional *CEFR*-aligned corpus of Dutch EFL argumentative writing. The aim was to assess the use of this feature of Dutch English in formulating a supplementary coherence and cohesion descriptor that might help to differentiate within the relatively under-defined expanse of *CEFR C2* for our own local context. The need for differentiation, which has been commented on by Martyniuk and Noijons (2007) and Springer (2012) among others, was highlighted by the fact that

nearly half of the cohort of first-year students of English Language and Culture that participated in this study was squarely placed at *CEFR C2* by the Online Oxford Placement Test (OOPT), apparently suggesting that the entry level in year 1 is the same as the projected exit level after three years of rigorous training.

As a preliminary, the observed differences between Dutch learners and native speakers of English were tested for statistical significance, confirming that the mean number of local anchors per main clause in Dutch English novice writing is significantly higher than in NS novice writing of the same genre. What is more, learners' proficiency, operationalised as OOPT score, was found to be weakly but significantly inversely correlated with their use of local anchors, suggesting that use of local anchors is one of the subtle features that may help to define language development at higher levels of acquisition. The proposed supplementary descriptor serves to capture this development and to adapt the *CEFR* to the local context of (future) Dutch English language professionals.

7.1.3

Chapter 4 presented a comprehensive longitudinal Bayesian analysis of initial adverbial use that was intended to address the discrepancies in text types, reference corpora, normalisation and adverbial classification between previous studies and make up for the lack of longitudinal statistical testing. It also extended the scope of previous studies by considering a) use of initial adverbials overall, b) use of initial circumstance, stance and linking adverbials and c) use of local anchors in 780,414 words of syntactically-parsed text in total, comprising longitudinal data from four cohorts of Dutch students of English Language and Culture as well as two NS reference corpora. For each of the relevant measures, we did not only contrast Dutch learners and NSs, but also investigated to what extent the frequency with which adverbials are used initially depends on text type, academic

discipline, longitudinal development and whether or not written assignments were timed or untimed.

The analysis revealed that in the NS and NNS corpora alike, linking adverbials predominate in the argumentative essays, while circumstance adverbials are used more often in the literature essays and articles. For each of the text types considered, learners use more initial adverbials than NSs. More particularly, learners use more circumstance and linking adverbials, but fewer stance adverbials than both novice and expert native writers. In any adverbial class, they also use more adverbials for local anchoring, apparently transferring Dutch discourse linking strategies to their L2 English. In the course of their studies, learners generally develop in the direction of native writing, but the gap between learners and NSs in the use of linking adverbials either persists or widens, depending on text type. This suggests that the potential role of teaching warrants further research. For each of the adverbial types considered, untimed essays are more native-like than timed essays. As the difference between year 1 timed and untimed argumentative essays is in many cases bigger than the difference between year 1 and year 2 argumentative essays, it appears that the effect of increased knowledge of or exposure to English interacts with students' reliance on default strategies that they fall back on under time pressure.

7.1.4

Chapter 5 turned to two (possibly interrelated) causes of Dutch learners' heavy reliance on initial adverbials to achieve textual cohesion: transfer and teaching. It started out by investigating the use of initial adverbials in a corpus of Dutch novice writers' argumentative essays written in their L1 Dutch and L2 English as well as a corpus of argumentative essays written by L1 English novice writers. This corpus analysis revealed L1-interlanguage congruity in Dutch learners' use of initial

linking adverbials and local anchors in their L1 and L2, thereby satisfying one of Jarvis' (2000) criteria for establishing transfer. Dutch learners turned out to rely even more heavily on initial linking adverbials and adverbials used for local anchoring in their L1 Dutch than in their L2 English. The incidence of initial linking adverbials and local anchors in the L1 English argumentative essays was considerably lower than in either the L1 Dutch or L2 English texts. However, a transfer-based account fails to explain why frequent use of other types of initial adverbials in L1 Dutch, most notably those that fall into the class of stance adverbials, is not reflected to a similar extent in Dutch learners' L2 English.

Frequent use of linking devices by other learner populations has previously been ascribed to teaching (cf. Crewe, 1990; Field & Yip, 1992; Milton & Tsang, 1993; Leńko-Szymańska, 2008). The possibility that teaching might also play a role in Dutch learners' characteristic use of initial linking adverbials was therefore investigated by examining their frequency of occurrence and treatment in L1 Dutch and L2 English secondary-school coursebooks and final secondary-school reading exams. Treatment of linking adverbials in the L2 English coursebooks appeared to depend on the amount of writing practice offered and varied from presentation in (contextless) lists categorised according to semantic function to frequent reminders to use linking adverbials for text structuring. L1 Dutch coursebooks pay even more attention to linking devices and focus heavily on establishing logical relations between sentences by paying attention to what they mostly refer to as 'signalling words'. While the final secondary-school Dutch reading exam has a similar emphasis on structure, logical relations and argumentative functions, the final English reading exam includes relatively few questions about linking devices. Nevertheless, they feature heavily in dedicated exam preparation, possibly because learners are felt to benefit from a focus on linking devices to keep track of the argumentation of exam texts, especially if the

demonstrable presence of low-frequency vocabulary hampers understanding. It appears that Dutch learners' possibly transfer-induced reliance on initial adverbials for discourse linking may be reinforced by a largely reductionist approach to teaching textual cohesion, leading learners to believe that repeated use of linking words will help them achieve instant cohesion.

7.1.5

Chapter 6 further explored the validity of a V2 transfer-based explanation of advanced Dutch learners' use of initial adverbials by investigating the extent to which the specific pattern of initial adverbial use found in earlier studies is unique to Dutch learners of English. It presented a quantitative and qualitative analysis of initial adverbial use in argumentative novice writing by NSs and advanced learners of English whose L1 is either Dutch (Germanic V2) or French (Romance SVO) in order to determine whether Dutch learners' tendency to place adverbials in pre-subject position is a) a transfer-induced feature of Dutch English b) a more widely-shared interlanguage feature, or c) characteristic of novice writing in general. Our quantitative analysis led us to conclude that Dutch learners of English are not exceptional in their relatively frequent use of adverbials in pre-subject position. In fact, Francophone learners appear to use even more, while the NS novice writers use considerably fewer initial adverbials than either learner group. This would suggest that frequent use of initial adverbials is neither a V2 transfer effect nor a defining characteristic of novice writing in general. Rather, it appears to be an interlanguage feature shared by Francophone and Dutch EFL learners alike.

A qualitative analysis of the use of these adverbials in context, however, reveals differences not only between learners and NSs but also between both learner groups. Dutch learners have a stronger tendency than either other group to use initial adverbials for local anchoring, apparently transferring the discourse

linking function of the preverbal slot in their L1 Dutch. Francophone learners' use of initial adverbials, on the other hand, often appears to be motivated by stylistic preferences – rather than structural differences – transferred from their L1. Considerations of style, articulated in French style guides, may play a role in Francophone learners' frequent use of initial adverbials in their L2 English as well. Writers may opt for placement in initial position to avoid consecutive sentences starting with a subject, to create a rhythmic effect, similar to written French, by coordinating strings of initial adverbial phrases or to achieve balance within the sentence, particularly if it has a relatively light subject. It appears, therefore, that EFL writing by different L1 groups may ostensibly be the same in writers' heavy reliance on adverbials in initial position, but at the same time markedly different in the way these adverbials are realised and the pragmatic purposes they serve.

7.2 Implications

The accumulated evidence suggests that it is not so much the overall frequency of initial adverbials that sets apart advanced Dutch learners' EFL writing from the writing of novice and expert NSs and EFL learners with other L1 backgrounds, but the way initial adverbials are used for discourse linking purposes. After all, in terms of overall frequency the Dutch learners are outstripped by the French. Also, the differences between Dutch learners and NS writers that are consistent across text-types and cohorts and that are persistent between year 1 and 3 are Dutch learners' frequent – and in some cases increasing – use of initial linking adverbials and their use of initial adverbials (regardless of adverbial class) for local anchoring, both of which have a cohesive function.

While a contrastive analysis like this may reveal differences between language learners and native speakers, it does not automatically follow that these differences should be interpreted as deficits (House & Kasper, 2000, qtd. in Barron,

2011). The inclusion of novice and expert NS texts, texts belonging to different genres and academic fields as well as L2 texts produced by writers with different L1 backgrounds has revealed patterns of adverbial use that characterise different ‘discourse accents’ (Scarcella, 1983). The use of clause-initial adverbials for discourse linking purposes therefore appears to be a feature of what might be called a Dutch English discourse accent.

Two separate underlying causes for this linguistic behaviour may be distinguished. Subtle traces of transfer at the syntax-pragmatics interface are likely to lie at the root of advanced Dutch learners’ use of initial adverbials for local anchoring. The corpus analyses reported on in this thesis have provided qualified evidence for three types of effects that would support a transfer-based account: a) intra-L1-group homogeneity, b) inter-L1-group heterogeneity and c) L1-interlanguage congruity (Jarvis 2000). Intra-L1-group homogeneity is evident in the consistency – across cohorts and text types – with which advanced Dutch learners use initial adverbials for local anchoring. Inter-L1-group heterogeneity can be observed in the divergence between advanced Dutch learners of English on the one hand and native speakers and Francophone learners of English on the other. L1-interlanguage congruity, finally, manifests itself in Dutch learners’ preference for local anchoring in their L1 Dutch as well as in their L2 English. It appears, in short, that local anchoring is a discourse linking strategy employed much more frequently and consistently by the L1 Dutch writers than by either other group, in their L1 Dutch as well as in their L2 English, all of which adds up to substantiate the hypothesised role of pragmatic-syntactic transfer.

At the same time, transfer does not fully account for the frequent use of initial linking adverbials as a feature of Dutch English, in spite of the L1-interlanguage congruity observed between Dutch learners’ L1 Dutch and L2 English texts. First of all, the considerable increase in initial linking adverbials in

Dutch learners' literature essays between year 1 and 3 suggests that their use may be guided by a conscious effort inspired by explicit instruction rather than inadvertent L1 influence. Another reason why a transfer-based account is problematic is that a significantly higher mean frequency – relative to the mean frequencies found in novice and expert NS texts – can be observed not only in L1 Dutch but also in L1 French EFL writing. The idea that heavy reliance on initial linking devices is a more widely shared interlanguage feature is consistent with studies into the writing skills of EFL learners with various L1 backgrounds, which have pointed towards the role of coursebooks in creating the impression that the more linking devices learners use, the more structured and coherent their text will be (cf. Crewe, 1990; Field & Yip, 1992; Milton & Tsang, 1993; Leńko-Szymańska, 2008).

In the Dutch context, this impression seems to be confirmed by the largely reductionist approach to teaching textual cohesion observed at secondary-school level, in which a focus on linking devices comes at the expense of representative discussion of other cohesive strategies in L2 English coursebooks. Meanwhile, the L1 Dutch curriculum displays an even heavier emphasis on logical relations and the 'signalling words' that serve as their signposts. Learners are not only encouraged to use them to flag each twist and turn in the argumentation of their own writing but also to rely on them as text analysis tools in the final reading exam. If Dutch learners' frequent use of initial linking adverbials in their L2 English mirrors their heavy reliance on this same strategy in their L1 Dutch, this may therefore not necessarily or not exclusively be the result of transfer but rather of conscious application of the writing conventions taught at secondary school both in their L1 Dutch and in their L2 English writing.

7.3 Future directions

The focus of this study has been on the frequency and pragmatic use of initial adverbials as a defining feature of advanced Dutch English. While a comparison between the various sub-corpora has yielded insight into the distributional patterns of initial adverbials in various discourse accents as well as the role of discipline, text type and L1 influence in these patterns of initial adverbial use, it cannot confirm whether the relatively high frequencies found for both groups of L2 writers result from a preference for placement in initial rather than medial or final position (cf. Granger & Tyson, 1996; Altenberg & Tapper, 1998; Gilquin *et al.*, 2007) or might reflect higher overall frequencies. The syntactic annotation procedure used in the present study, combined with automatic retrieval of adverbials through Corpus Studio (Komen, 2012), could very well be applied to adverbials in other sentence positions. This would serve to paint a more complete picture and shed light on the potential existence of positional preferences as an interlanguage universal (cf. Gilquin *et al.*, 2007).

While a Contrastive Interlanguage Analysis (Granger, 1996) may reveal features that characterise the writing of a particular group of L1 or L2 writers, this study has also illustrated that use of such features is not necessarily stable over time. In order to trace linguistic development, the use of longitudinal corpora such as LONGDALE (Meunier, 2015) is clearly invaluable. The use of such corpora could be further enhanced through calibration with the *Common European Framework of Reference* to ensure comparability of different L1 sub-corpora. The limited reliability of using institutional status as an indication of proficiency level means, for example, that the comparison of L1 Dutch and L1 French learners reported on in chapter 6 may potentially be slightly uneven, which could have been avoided or taken into account by means of a more objective measure of both groups' proficiency level. A further benefit of using a *CEFR*-aligned corpus is that

it facilitates a comparison between more and less proficient users within one sub-corpus. Chapter 3 described the use of a cross-sectional *CEFR*-aligned sub-corpus of one cohort of Dutch students of English Language and Culture to determine whether the incidence of local anchoring could be seen as a quantifiable determinant for language development at higher stages of acquisition. The fact that this cohort has now completed their Bachelor degrees opens up the exciting possibility of exploring a proficiency-stratified longitudinal sub-corpus of student writing. It would be particularly interesting to see what patterns emerge from tracing individual users' development over time in relation to their objectively measured proficiency scores on an Online Oxford Placement Test administered at the beginning and end of their studies.

Another possible improvement and direction for future research regards the selection of reference corpora. The inclusion of an L1 English novice as well as expert corpus has facilitated a 'fair' comparison between L1 and L2 English novice writing in addition to the comparison between L2 English novice writing and the kind of L1 expert writing most frequently held up as a model of good academic writing. This serves to demonstrate that frequent use of initial (linking) adverbials is not necessarily a feature of novice writing in general, as there are demonstrable differences between L1 and L2 novice writing. However, differences in text type between the L1 English corpora prevented a direct comparison of L1 novice and expert writing as well as a threeway comparison of L1 and L2 novice writing and L1 expert writing, which would have allowed a more direct investigation into the role of writing expertise vs L1 influence.

Similarly, chapter 5 reported on the L1-interlanguage congruity of Dutch learners' use of linking adverbials, which points to the possible role of transfer, but also noted that a heavy emphasis on signposting logical relations in L1 Dutch coursebooks might reinforce learners' reliance on linking adverbials as a default

discourse linking strategy. By definition, a comparison of L1 Dutch, L1 English and L2 English novice writing alone cannot comment on the degree to which these factors are intertwined. While the frequency of initial linking adverbials in Dutch novice writers' L1 Dutch and L2 English texts may both to some extent be a teaching effect, the considerable importance attached to linking adverbials in the L1 Dutch curriculum may in turn be a natural reflection of the fact that Dutch, as a V2 language, can easily accommodate different types of constituents, including adverbials, in the preverbal slot and tends to use these for discourse linking purposes. In this respect, it would be interesting to compare the corpus of L1 Dutch novice writing used in the present study with a corpus of L1 Dutch expert writing, to disentangle the respective roles of the teaching of writing conventions and the transfer of discourse linking preferences that may be inherent in the language.

Another interesting question is to what extent a strong emphasis on linking adverbials as signposts for logical relations and text structuring in the L1 Dutch secondary-school curriculum may represent a washback effect from the nationwide final reading exam, which relies heavily on these signposts as text-analysis tools. As 'tests are held to be powerful determinants of what happens in classrooms' (Alderson, 1993, p. 115), the issue of test validity of standardised L1 as well as L2 final reading exams clearly deserves close scrutiny. Anecdotal evidence suggests, for example, that systematic training in the recognition and interpretation of signalling words and other reading strategies allows students to do well in L2 English reading exams, to the extent that they reportedly outperform educated, but untrained, native speakers. Kwakernaak and Pouw (1999) also voice many teachers' concerns that multiple-choice reading-test items may test learners' mastery of a range of reading strategies which have limited relevance beyond learners' school careers, rather than true understanding of an L2 text. Similarly, in an experiment reported by Tillema (2006), the limited discrepancy between Dutch

learners' scores on an L2 English final reading exam and a Dutch translation of the same exam suggested that final reading exams might lack L2-specificity. While this does not in any way prove whether or not learners may benefit from a focus on reading strategies in general and signalling words in particular, it is well worth investigating what secondary-school L2 reading exams actually test and to what extent performance on test items may be predicted by textual features, such as the lexical profile of texts and questions or overlap between texts and multiple-choice options, as well as learner features, such as vocabulary size or general proficiency level (cf. Ozuru *et al.*, 2008; Nergis, 2013).

In spite of the observed difference between Dutch learners and NSs, the use of local anchors in English is not grammatically wrong. Acceptability judgments by native speakers are therefore not likely to pick up on these subtle traces of transfer. However, if local anchors, which present discourse-old information in what in present-day English is an information-structurally prominent position, can be considered the written equivalent of 'superfluous accents' in spoken language (Dimitrova, 2012), they may affect ease of discourse integration. This could possibly be measured by means of EEG, helping us to determine whether the differences between advanced learners and L1 English writers described in this thesis reflect cognitive differences in the processing of adverbial-antecedent pairs in clause-initial position. In this way, a combination of corpus-based and experimental methods could lead us from a descriptive account of the differences between L1 and L2 discourse accents to the neuro-linguistic reality underlying these differences.

A more practical question is whether the Dutch English features identified in this collection of studies should be targeted for pedagogical intervention. As Granger (2009) has argued, this depends on a number of factors, 'including learner needs, teaching objectives and teachability' (qtd. in Guilquin *et*

al., 2007). It is clear that the learner needs of a population of future language professionals, i.e. our Dutch students of English Language and Culture, go over and beyond that of most other language learners (De Haan & Van der Haagen, 2012; Flowerdew, 2015). It is equally clear that where heavy reliance on initial linking adverbials is a symptom of 'surface logicity' (Crewe, 1990), a shortcut to instant cohesion, this is more than just a linguistic feature and deserves instructors' attention. The communicative effect of use of initial adverbials for local anchoring is less obvious. The use of a focus position for local anchors intended as neutral discourse links may be marked but is not likely to seriously impede effective communication. It would therefore be difficult to defend a prescriptive ban on local anchoring. However, in order to distinguish themselves in a largely multilingual society, in which a considerable proportion of the population has an adequate command of English, students of English Language and Culture can undoubtedly benefit from an awareness of features of different varieties of English, including ways of achieving cohesion and highlighting different elements in the sentence.

In addition to existing L1 corpus-informed teaching materials, recent years have seen the introduction of learner corpus-informed teaching materials, such as *Macmillan's English Dictionary for advanced learners* (cf. Gilquin *et al.*, 2007), which go a long way towards helping learners become aware of more widely shared interlanguage features. Similar L1-specific resources are hardly feasible for publishers, however. In fact, doing justice to the considerable variation across disciplines, such as between the NS linguistics and literature articles that were compared in chapter 4, would require not only L1-specific but also discipline-specific materials. Unlike published materials, data-driven learning methods can accommodate both the L1-specific and discipline-specific local context and present an opportunity to use corpus data to raise advanced,

specialised learners' awareness of variation along these different dimensions. Harwood (2005) has suggested that 'corpus data may be used as a launch pad for classroom research into how the linguistic item in question is used by experts and students in the learners' local context' (p. 158). Other advantages, cited by Boulton (2009), include fostering learner autonomy, increasing language awareness and improving ability to deal with 'the complexity and fuzziness of authentic data of a foreign or second language' (pp. 2-3). For future language professionals in particular, turning learners into researchers of their own and others' language use seems a highly appropriate methodological choice. In this way, learners do not only have to be subject to other people's scrutiny; they can be applied linguists in their own right.

References

- Alderson, J. & Wall, D. (1993). Does washback exist? *Applied linguistics*, 14(2), 115-129.
- Alderson, J. (2007). The *CEFR* and the Need for More Research. *The Modern Language Journal*, 91(4), 659-663.
- Altenberg, B. & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80-93). Harlow: Longman.
- Bardovi-Harlig, K. (1999). Exploring the interlanguage of interlanguage pragmatics: A research agenda for acquisitional pragmatics. *Language learning*, 49(4), 677-713.
- Barends, T., Van Biemen, J., Ekkelboom, S., Grezel, J., Van Hoek, G., Jansen, I., ... Taihutu, T. (2009). *Nieuw Nederlands – Samenvatting onderbouw havo/vwo*. Groningen: Noordhoff.
- Barron, A. (2012). Interlanguage pragmatics: From use to acquisition to second language pedagogy. *Language Teaching*, 45(01), 44-63.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finnegan, E. (1999). *The Longman grammar of spoken and written English*. Harlow: Longman.
- Bohnacker, U., & Rosén, C. (2007). Transferring information-structural patterns from Swedish to German. In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 27-38). Somerville, MA: Cascadilla Proceedings Project.

- Bohnacker U. & Rosén, C. (2008). The clause-initial position in L2 German declaratives: Transfer of information structure. *Studies in Second Language Acquisition*, 30(4), 511-538.
- Bouma, K. (2016, August 26). Meer dan de helft van de studies volledig in het Engels. *De Volkskrant*. Retrieved January 15, 2017, from <http://www.volkskrant.nl/binnenland/meer-dan-de-helft-van-de-studies-volledig-in-het-engels~a4364526/>
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(01), 37-54.
- Callies, M. (2007). Extending the scope of inquiry in interlanguage pragmatics: The case of focus constructions. *Crosscultural pragmatics and interlanguage English*, 11, 73-90.
- Callies, M. (2009). *Information highlighting in advanced learner English: The syntax-pragmatics interface in second language acquisition*. Amsterdam & Philadelphia: John Benjamins.
- Carroll, M. & Lambert, M. (2003). Information structure in narratives and the role of grammaticised knowledge: A study of adult French and German learners of English. In C. Dimroth & M. Starren (Eds.), *Information structure and the dynamics of language acquisition* (pp. 267-87). Amsterdam: Benjamins.
- Carroll, M. & Lambert, M. (2006). Reorganizing principles of information structure in advanced L2s. *Educating for advanced foreign language capacities: Constructs, curriculum, instruction, assessment*, 54-73.
- Carroll, M., Murcia-Serra, J., Watorek, M. & Bendiscioli, A. (2000). The relevance of information organization to second language acquisition studies: The descriptive discourse of advanced adult learners of German. *Studies in Second Language Acquisition*, 22(3), 441-66.

- Charolles, M. (2003). De la topicalité des adverbiaux détachés en tête de phrase. *Travaux de Linguistique*, 47(2), 11-51.
- Chen, C. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1), 113-130.
- Cobb, T. (n.d.) *Range for texts* (version 3) [software]. Available from <http://www.lex Tutor.ca/cgi-bin/range/texts/index.pl>.
- Council of Europe. (1992). *Transparency and coherence in language learning in Europe: objectives, evaluation, certification*. Report on the Rüscliikon Symposium. Strasbourg: Council of Europe.
- Council of Europe. (2001). *Common European Framework of Reference for: Learning, Teaching, Assessment*. Retrieved February 25, 2015, from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf.
- Crewe, W. (1990). The illogic of logical connectives. *ELT Journal*, 44(4), 316-325.
- De Haan, P. & Van der Haagen, M. (2012). Modification of Adjectives in Very Advanced Dutch EFL Writing: A Development Study. *The European Journal of Applied Linguistics and TEFL*, 1(1), 129-142.
- Deloffre, F. (1967). *La phrase Française*. Paris: Société d'Édition d'Enseignement Supérieur.
- De Mönning, I. (2000). Parsing a learner corpus? *Language and Computers*, 33, 81-90.
- Dimitrova, D., Stowe, L., Redeker, G., & Hoeks, J. (2012). Less is not more: Neural responses to missing and superfluous accents in context. *Journal of cognitive neuroscience*, 24(12), 2400-2418.
- Dimroth, C., & Starren, M. (Eds.). (2003). *Information structure and the dynamics of language acquisition* (Vol. 26). John Benjamins.

- Dreschler, G. & Hebing, R. (2010). Sentence beginnings in Present-Day English, Present-Day Dutch and Old English. In C. Cummins, C. Elder, T. Godard, M. Macleod, E. Schmidt and G. Walkden (Eds). *Proceedings of the Sixth Cambridge Postgraduate Conference in Language Research* (pp. 58-69). Cambridge, UK: Cambridge Institute of Language Research.
- Dreschler, G. (2015). *Passives and the Loss of Verb Second: A study of Syntactic and Information-Structural Factors*. Utrecht: LOT.
- Edwards, A. (2016). *English in the Netherlands: Functions, forms and attitudes*. Amsterdam: John Benjamins.
- EF English Proficiency Index. (2015). Retrieved January 15, 2016, from <http://media2.ef.com/~/media/centralescom/epi/downloads/full-reports/v5/ef-epi-2015-english.pdf>.
- ERK. (n.d.) *Streefniveaus HAVO/VWO*. Retrieved July 1, 2015, from <http://www.erk.nl/docent/streefniveaus/havo/>.
- Eurobarometer. (2012). *Europeans and their Languages*. Conducted by TNS Opinion & Social at the request of Directorate-General Education and Culture, Directorate-General for Translation and Directorate-General for Interpretation. Survey co-ordinated by the European Commission, Directorate-General for Communication (DG COMM Research and Speechwriting Unit).
- Ferguson, G., Pérez-Llantada, C. & Plo, R. (2011). English as an international language of scientific publication: A study of attitudes. *World Englishes*, 30(1), 41-59.
- Field, Y. & Yip, L. (1992). A comparison of internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal*, 23(1), 15-28.

- Figueras, N. (2007). The *CEFR*, a Lever for the Improvement of Language Professionals in Europe. *The Modern Language Journal*, 91(4), 673-675.
- Flowerdew, J. (2015). Some thoughts on English for research publication purposes (ERPP) and related issues. *Language Teaching*, 48(2), 250-262.
- Fuchs, C. & Fournier, N. (2003). Du rôle cadratif des compléments localisants initiaux selon la position du sujet nominal. *Travaux de Linguistique*, 47(2), 79-09.
- Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly*, 1(4), 253-266.
- Gilquin, G., Granger, S. & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, 6(4), 319-335.
- Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg & M. Johansson (Eds.), *Languages in Contrast. Text-based Cross-linguistic Studies* (pp. 37-51). Lund: Lund University Press.
- Granger, S. & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27.
- Granger, S. (2004). Computer learner corpus research: current status and future prospects. *Language and Computers*, 52(1), 123-145.
- Grevisse, M. & Goosse, A. (1993). *Le Bon Usage*. Paris: Duculot.
- Halliday, M. & Hassan, R. (1976). *Cohesion in English*. London: Longman.
- Hannay, M. & Keizer, M. (1993). Translation and contrastive grammar: The grammatical vs the communicative strategy. *Toegepaste taalwetenschap in artikelen*, 45(1), 65-88.
- Hannay, M., & Mackenzie, J. L. (2002). *Effective writing in English: a sourcebook*. Bussum: Coutinho.

- Hartog, S., Huizing, E., Marinus, N., Meeuwssen, N., Van Ringen, J., Sewnandan, N., ... & Wijnhoven, M. (2008). *Stepping Stones: 2e fase. Leerjaar 4. VWO, Reference*. Groningen: Noordhoff.
- Hartog, S., Marinus, N., Meeuwssen, N., van Ringen, J., Sewnandan, N., Tillotson, P., ... & Wijnhoven, M. (2009). *Stepping Stones: 2e fase. Leerjaar 5. VWO, katern 2*. Groningen: Noordhoff.
- Harwood, N. (2005). What do we want EAP teaching materials for? *Journal of English for Academic Purposes*, 4(2), 149-161.
- Hasselgård, H. (2009). Thematic choice and expressions of stance in English argumentative texts by Norwegian learners. In K. Aijmer (Ed.), *Corpora and Language Teaching* (pp. 121-140). Amsterdam: Rodopi.
- Hawkins, J. & Filipovic, L. (2012). *Criterion features in L2 English: Specifying the Reference Levels of the Common European Framework* (Vol. 1). Cambridge: UCLES/Cambridge University Press.
- Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, 37(2), 275-301.
- Hoek, P. (2012). *Een moeilijke tekst of een moeilijke vraag?* Unpublished MA thesis. University of Amsterdam.
- Honders, H., Jongejans, L., Budding-Derks, J. & Van Haaren, S. (2014). *Examentraining: Vwo Engels 2015*. Zeist: L & M educatief.
- Hulstijn, J. (2007). The Shaky Ground beneath the *CEFR*: Quantitative and Qualitative Dimensions of Language Proficiency. *The Modern Language Journal*, 91(4), 663-667.
- Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, 50(2), 245-309.

- Jongsma, H., Flier, E., Merckx, P. & Van de Pol, J. (2010). *Op Niveau Tweede Fase Verwerkingsboek 4 VWO*. Utrecht: ThiemeMeulenhoff.
- Kirkwood, H. (1969). Aspects of word order and its communicative function in English and German. *Journal of Linguistics*, 5(1), 85-107.
- Klein, D. & Manning, C. (2003). Accurate Unlexicalized Parsing. *Proceedings of the 41st Meeting of the Association for Computational Linguistics* (pp. 423-430). Sapporo, Japan: Association for Computational Linguistics.
- Komen, E. (2012). Coreferenced corpora for information structure research. In J. Tyrkkö, M. Kilpiö, T. Nevalainen & M. Rissanen (Eds.), *Outposts of Historical Corpus Linguistics: From the Helsinki Corpus to a Proliferation of Resources*. Helsinki: Research Unit for Variation, Contacts, and Change in English. Retrieved March 20, 2013, from <http://www.helsinki.fi/varieng/journal/volumes/10/komen/>.
- Komen, E. (2013). *Finding Focus: a study of the historical development of focus in English*. Utrecht: LOT.
- Kruschke, J. (2010). *Doing Bayesian data analysis: A tutorial introduction with R*. Academic Press.
- Kwakernaak, E. & Pouw, A. (2000). Toetsing en vernieuwing in het vreemdetalenonderwijs. *Levende Talen Tijdschrift*, 1(1), 44-51.
- Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- Laskin, L. (2016). *Frequentist And Bayesian Approaches In Statistics*. Retrieved June 17, 2016, from <http://www.probabilisticworld.com/frequentist-bayesian-approaches-inferential-statistics/#more-1677>.
- Laufer, B. & Nation, P. (1999). A vocabulary size test of controlled productive ability. *Language Testing*, 16(1), 33-51.

- Lemmens, M. & Parr, T. (1995). *Handboek voor de vertaler Nederlands-Engels*. Groningen: Wolters-Noordhoff.
- Leńko-Szymańska, A. (2008). Non-native or non-expert? The use of connectors in native and foreign language learners' texts. *Acquisition et interaction en langue étrangère*, (27), 91-108.
- Little, D. (2007). The Common European Framework of Reference for Languages: Perspectives on the Making of Supranational Language Education Policy. *The Modern Language Journal*, 91(4), 645-655.
- Litré, D. (2014). *A Cognitive, Longitudinal Study of the Use of the English Present Progressive by Intermediate and Advanced French-Speaking Learners* (unpublished doctoral dissertation). Université Catholique de Louvain, Louvain.
- LOCNESS. (2005). Retrieved June 24, 2011, from <http://www.corpus4u.org/forum/upload//forum/2005061503142048.pdf>.
- Los, B. (2009). The loss of verb-second and the switch from bounded to unbounded systems. In A. Meurman-Solin, M. J. López-Couso & B. Los (Eds.), *Information structure and syntactic change in the history of English* (pp. 21-46). Oxford: Oxford University Press.
- Los, B. & Dreschler, G. (2012). The loss of local anchoring: From adverbial local anchors to permissive subjects. In T. Nevalainen & E. Closs Traugott (Eds.), *The Oxford handbook of the history of English* (pp. 859-872). Oxford: Oxford University Press.
- Magnus, I. (2014). The distribution of adverbials in declarative sentences in French. In N. Delbecque, K. Lahousse & W. van Langendonck (Eds.), *Non-nuclear Cases* (pp. 173-216). Amsterdam/Philadelphia: John Benjamins.

- Martyniuk, W. & Noijons, J. (2007). *The executive summary of results of a survey on the use of the CEFR at national level in the Council of Europe Member States*. Retrieved February 16, 2015, from http://www.cito.nl/onderzoek%20en%20wetenschap/achtergrondinformatie/int_onderzoek/CEFR_publicaties_engels.
- McNamara, T. (2007). Language Assessment in Foreign Language Education: The Struggle over Constructs. *The Modern Language Journal*, 91(2), 280-282.
- Meestringa, T. & Ravesloot, C. (2013). Schrijven in de tweede fase. *Levende Talen Magazine*, 6, 6-10.
- Meunier, F. (2015). Introduction to the LONGDALE project. In E. Castello, K. Ackerley, F. Coccetta (Eds.), *Studies in Learner Corpus Linguistics: Research and Applications for Foreign Language Teaching and Assessment*. Bern: Peter Lang.
- Milton, J. & Tsang, E. (1993). A corpus-based study of logical connectors in EFL students' writing: directions for future research. In: *Studies in Lexis. Proceedings of a seminar on lexis organized by the Language Centre of the HKUST* (pp. 215-246). Hong Kong: HKUST.
- Mittwoch, A., Huddleston, R. & Collins, P. (2002). The clause: adjuncts. In R. Huddleston & G. Pullum (Eds.), *The Cambridge grammar of the English language* (pp. 663-784). Cambridge, UK: Cambridge University Press.
- Mulder, E., Den Tenter, P., De Boer, T., Dirksen, J. & Van Son, J. (2014). *Talent. Nederlands voor de onderbouw handboek 2 havo/vwo*. 's-Hertogenbosch: Malmberg.
- Newbrook, J. & Wilson, J. (2003). *Proficiency Gold*. Harlow: Longman.
- Nergis, A. (2013). Exploring the factors that affect reading comprehension of EAP learners. *Journal of English for Academic Purposes*, 12(1), 1-9.
- Norris, J. & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis

- and quantitative meta-analysis. *Language Learning*, 50(3), 417-528.
- Ozuru, Y., Rowe, M., O'Reilly, T., & McNamara, D. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods*, 40(4), 1001-1015.
- Pols, G. (2016, May 11). Loonkloof iets kleiner, vrouwen verdienen nog altijd minder. *De Volkskrant*. Retrieved January 15, 2017, from <http://www.volkskrant.nl/economie/loonkloof-iets-kleiner-vrouwen-verdienen-nog-altijd-minder~a4299115/>
- Purpura, J. (n.d.). *The Oxford Online Placement Test: What Does It Measure and How?* Retrieved June 30, 2015, from <https://www.oxfordenglishtesting.com/defaultmr.aspx?id=3048>.
- Putten, C. & Verploegh, H. (2015). *Examenbundel VWO Engels: Jouw beste voorbereiding op je examen in 2016*. Amersfoort: ThiemeMeulenhoff.
- Rosén, C. (2006). "Warum klingt das nicht deutsch?": *Probleme der Informationsstrukturierung in deutschen Texten schwedischer Schüler und Studenten* (Vol. 67). Almqvist & Wiksell International.
- Rundell, M. (Ed.). (2007). *Macmillan English dictionary for advanced learners* (2nd ed.). Oxford: Macmillan Education.
- Scarcella, R. (1983). Discourse accent in second language performance. In S. Gass & L. Selinker (Eds.), *Language Transfer in Language Learning* (pp. 306-326). Rowling, MA: Newbury House.
- Schmitt, N., Jiang, X. & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43.
- Shaw, P. (2004). Sentence openings in academic economics articles in English and Danish. *Nordic Journal of English Studies*, 3(2), 67-84.

- Soars, L., Soars, J., & Wheeldon, S. (2005). *New Headway: Upper-Intermediate*. Oxford: Oxford University Press.
- Springer, P. (2012). *Advanced learner writing: A corpus-based study of the discourse competence of Dutch writers of English in the light of the C1/C2 levels of the CEFR*. Oisterwijk: BOXpress.
- Target Training. (n.d.) *Business English Training and the CEFR*. Retrieved February 25, 2015, from <http://www.targettraining.eu/business-english-and-the-CEFR/>.
- Thomas, A. & Burgess, S. (2014). *Gold: Advanced*. Harlow: Pearson Longman.
- Tillema, M. (2006). CITO leestoetsen voor VWO: T2-specificiteit en scharnierwoorden. Unpublished MA thesis. Utrecht University.
- Van Noord, G. (2006). At Last Parsing Is Now Operational. In *TALN06. Verbum Ex Machina. Actes de la 13e conference sur le traitement automatique des langues naturelles* (pp. 20-42). Louvain, Belgium: Presses universitaires de Louvain.
- Verheijen, L., Los, B. & De Haan, P. (2013). Information structure: The final hurdle? The development of syntactic structures in (very) advanced Dutch EFL writing. *Dutch Journal of Applied Linguistics*, 2(1), 92-107.
- Virtanen, T. (2004). *Approaches to cognition through text and discourse*. Berlin: de Gruyter.

Appendix I

Table 1. Circumstance adverbials in *New Headway*, *Gold* and *Stepping Stones*

Circumstance	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
	upper intermediate	advanced	advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
after (that)							
after a while							
afterwards							
all day long							
at first							
at last							
because of							
by this time							
clearly							
due to							
especially							
eventually							
every February							
finally							
for the next 24 hours							
for this purpose							
frequently							
from early childhood							

232 - Traces of Transfer

Circumstance	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
	upper intermediate	advanced	advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
from an early age							
headfirst							
immediately							
in order (not) to							
in the same way							
in this way							
in three places							
in view of							
in winter							
just at that moment							
likewise							
next							
now							
one day							
owing to							
similarly							
suddenly							
then							
throughout his life							
today							
tomorrow							
two/several years ago							
up to that time							

	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
Circumstance	upper intermediate	advanced	advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
with great confidence							
with my family							
yesterday							

Table 2. Stance adverbials in *New Headway*, *Gold* and *Stepping Stones*

Stance	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
	upper intermediate	advanced	Advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
(un)fortunately							
above all							
actually							
admittedly							
annoyingly							
apparently							
as a matter of fact							
at least							
basically							
better still							
certainly							
clearly							
curiously							
definitely							
essentially							
even more importantly							
frankly							
funnily enough							
generally							
generally speaking							
happily							
honestly							

Stance	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
	upper intermediate	advanced	Advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
hopefully							
ideally							
in fact							
in my opinion							
in my view							
in reality							
in short							
in the main							
ironically							
irritatingly							
luckily							
maybe							
naturally							
no doubt							
not surprisingly							
obviously							
oddly enough							
of course							
on the whole							
personally							
presumably							
primarily							
probably							

236 - *Traces of Transfer*

Stance	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
	upper intermediate	advanced	Advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
quite honestly							
really							
sadly							
seriously though							
significantly							
strictly speaking							
surely							
surprisingly							
thankfully							
understandably							
undoubtedly							
unfortunately							

Table 3. Linking adverbials in *New Headway*, *Gold* and *Stepping Stones*

Linking	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
	upper intermediate	advanced	Advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
above all							
accordingly							
additionally							
after all							
all in all							
all the same							
also							
alternatively							
anyhow							
anyway							
apart from							
as a result							
as a result/consequence							
as well as							
as well as this							
at the end of the day							
basically							
besides							
besides (this)							
by the way							
by/in comparison							
by/in contrast							

	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
Linking	upper intermediate	advanced	Advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
consequently							
conversely							
despite (this)							
even so							
finally							
first of all							
firstly							
for a start							
for example							
for instance							
for this reason							
furthermore							
hence							
however							
in addition							
in addition to this							
in conclusion							
in contrast							
in other words							
in short							
in spite of							
in spite of/despite this							
in the first place							

	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
Linking	upper intermediate	advanced	Advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
in the second place							
incidentally							
indeed							
instead							
last							
last but not least							
moreover							
nevertheless							
next							
nonetheless							
of course							
on balance							
on the contrary							
on the one hand							
on the other hand							
on top of that							
overall							
rather							
rather the reverse							
secondly							
so							
still							
therefore							

240 - *Traces of Transfer*

	<i>New Headway</i>		<i>Gold</i>		<i>Stepping Stones</i>		
Linking	upper intermediate	advanced	Advanced	proficiency	Vwo 4 Kat 3	Vwo 4 Ref	Vwo 5 Kat 2
therefore, thus							
thirdly							
thus							
to sum up							
ultimately							
well							
yet							

Appendix II

Table 1. Adverbial categories included in the class of circumstance adverbials
(raw figures and proportional use)

	L1 Dutch	L1 French	NS
Addition/restriction	6 (2.49%)	7 (1.07%)	14 (6.14%)
Contingency	64 (26.56%)	74 (11.31%)	37 (16.23%)
<i>cause/reason</i>	24	42	21
<i>concession</i>	3	5	10
<i>condition</i>	26	19	2
<i>contrast</i>	2	4	2
<i>purpose</i>	9	2	1
<i>result</i>	0	2	1
Domain	14 (5.81%)	76 (11.62%)	20 (8.77%)
Extent/degree	0	1 (0.15%)	3 (1.32%)
Place	52 (21.58%)	152 (23.24%)	36 (15.79%)
Process	23 (9.54%)	45 (6.88%)	14 (6.14%)
Recipient	4 (1.66%)	1 (0.15%)	0
Time	78 (32.37%)	298 (45.57%)	104 (45.61%)
<i>Total</i>	241	654	228

Table 2. Adverbial categories included in the class of stance adverbials (raw figures and proportional use)

	L1 Dutch	L1 French	NS
Attitude	11 (12.64%)	39 (10.43%)	16 (15.38%)
Epistemic	76 (87.36%)	330 (88.24%)	81 (77.88%)
<i>Actuality and reality</i>	4	100	13
<i>Doubt and certainty</i>	27	97	37
<i>Imprecision</i>	0	10	6
<i>Limitation</i>	25	22	8
<i>Viewpoint or perspective</i>	19	97	16
<i>Source of knowledge</i>	1	4	1
Style	0 (0%)	5 (1.34%)	7 (6.73%)
<i>Total</i>	87	374	104

Table 3. Adverbial categories included in the class of linking adverbials (raw figures and proportional use)

	L1 Dutch	L1 French	NS
Addition	66 (19.76%)	155 (20.45%)	38 (16.89%)
Apposition	1 (0.3%)	14 (1.85%)	0
Contrast/concession	68 (20.36%)	254 (33.51%)	94 (41.78%)
Enumeration	81 (24.25%)	90 (11.87%)	21 (9.33%)
Exemplification	15 (4.49%)	84 (11.08%)	14 (6.22%)
Repetition	1 (0.3%)	6 (0.79%)	3 (1.33%)
Result/inference	77 (23.05%)	94 (12.40%)	43 (19.11%)
Summation	23 (6.89%)	44 (5.80%)	11 (4.89%)
Transition	2 (0.6%)	17 (2.24%)	1 (0.44%)
<i>Total</i>	334	758	225

Samenvatting (Summary in Dutch)

Vergevorderde Nederlandse leeders van het Engels, zoals studenten Engelse Taal en Cultuur aan de Radboud Universiteit, produceren over het algemeen goedlopende en doeltreffende Engelse teksten (vgl. Springer, 2012; De Haan & Van der Haagen, 2012). Toch vertoont het Engels van deze leeders nog herkenbaar Nederlandse invloeden. In tegenstelling tot de lagere niveaus van taalbeheersing, waar invloed van de moedertaal (T1) zichtbaar is in de doeltaal (T2) zowel op lexicaal als grammaticaal vlak, worden de hogere niveaus gekenmerkt door een meer subtiële vorm van interferentie op het gebied van pragmatiek (Bohnacker & Rosén, 2008; Callies, 2009). Dit is onder meer zichtbaar in de manier waarop leeders hun informatie structureren.

Er zijn aanwijzingen uit eerder verkennend onderzoek dat Nederlandse leeders verschillen van moedertaalsprekers in hun gebruik van initiële bijwoordelijke bepalingen (Verheijen *et al.*, 2013). Dit zou te maken kunnen hebben met een typologisch verschil tussen het Nederlands en het Engels. Het Engels is een *Subject-Verb-Object* (SVO) taal, waarin het onderwerp vooraf gaat aan het werkwoord. Eventuele bijwoordelijke bepalingen kunnen ofwel voor het onderwerp geplaatst worden, ofwel aan het einde van de zin. Het Nederlands, daarentegen, is een *Verb-second* (V2) taal, waarbij de persoonsvorm van het werkwoord in hoofdzinnen de tweede positie in de zin inneemt. De positie voorafgaand aan de persoonsvorm is zowel in syntactisch als informatie-structureel opzicht multifunctioneel en kan ingevuld worden met een onderwerp, een lijdend voorwerp of een bijwoordelijke bepaling (Los, 2009).

De eerste positie heeft een belangrijke verbindende functie. Vaak wordt die verbinding gemaakt met zinsverbindende bijwoordelijke bepalingen, zoals ‘toch’ in het voorbeeld hieronder:

1. Nog altijd verdienen Nederlandse vrouwen jaarlijks een bescheiden maandloon minder dan mannen. **Toch** neemt het verschil af en is de kloof vele malen kleiner dan in de meeste andere Europese landen.
(Pols, 2016)

Kenmerkend voor het Nederlands is ook dat zinnen vaak door middel van een bijwoordelijke bepaling met een expliciete verwijzing naar een antecedent in de voorgaande context worden ‘verankerd’ (Los en Dreschler, 2012). Voorbeelden zijn:

2. Het onderzoek laat niet alleen een positieve trend zien binnen Nederland, maar ook in verhouding met andere Europese landen. ‘We lijken naar de kopgroep te kruipen’, zegt Osse. **In die kopgroep** zitten verder vooral Scandinavische landen, en België.
(Pols, 2016)
3. Ik accepteer niet dat universiteiten puur en alleen uit commerciële overwegingen onderwijs in het Engels geven. **Hierover** is het laatste woord nog niet gezegd.
(Bouma, 2016)

De vetgedrukte bijwoordelijke bepalingen in de zinnen hierboven worden beide referentieel gebruikt. In het Engels is dit niet fout, maar wel gemarkeerd; voor een

neutrale verbinding met de voorgaande tekst wordt vaker het grammaticale onderwerp van de zin gebruikt.

Dit proefschrift onderzoekt in hoeverre het gebruik van initiële bijwoordelijke bepalingen kenmerkend is voor het geschreven Engels van vergevorderde Nederlandse leeders en hoe ze zich in dit opzicht ontwikkelen tussen het begin en het einde van hun studie:

1. Op welke manier verschillen vergevorderde leeders van moedertaalsprekers in de frequentie van initiële bijwoordelijke bepalingen en in de manier waarop ze die realiseren?
2. Hoe wordt het gebruik van initiële bijwoordelijke bepalingen door Nederlandse leeders van het Engels beïnvloed door a) hun niveau van taalbeheersing en b) hun ontwikkeling door de jaren heen?

De eerste twee onderzoeksvragen zijn gericht op het identificeren en beschrijven van het geschreven Engels van Nederlandstalige leeders; de overige zijn gericht op de rol van interferentie en van het onderwijs:

3. Welk bewijs is er voor de rol van interferentie in het gebruik van initiële bijwoordelijke bepalingen in het geschreven Engels van Nederlandse leeders?
 - a) In hoeverre zijn verschillende groepen Nederlandse leeders van het Engels vergelijkbaar in hun gebruik van initiële bijwoordelijke bepalingen in verschillende genres? Met andere woorden, is er sprake van intra-T1-groep homogeniteit?
 - b) Verschillen leeders van het Engels met een andere moedertaal van Nederlandstalige leeders in hun gebruik van initiële

bijwoordelijke bepalingen. Oftewel, is er sprake van inter-T1-groep heterogeniteit?

- c) Zijn er overeenkomsten tussen het gebruik van initiële bijwoordelijke bepalingen door Nederlandstalige leerders in hun eerste taal, Nederlands, en hun tweede taal, Engels? Dat wil zeggen, is er sprake van T1-T2 congruentie?

4. Speelt onderwijs een rol in het gebruik van initiële bijwoordelijke bepalingen door Nederlandse leerders?

Om deze vragen te beantwoorden is een contrastieve longitudinale corpusanalyse uitgevoerd van het geschreven Engels van studenten Engels tussen hun eerste en derde studiejaar, in vergelijking met dat van Franstalige leerders en moedertaalsprekers van het Engels. Deze teksten zijn automatisch syntactisch geparseerd, waarna initiële bijwoordelijke bepalingen zijn geïdentificeerd met behulp van Corpus Studio (Komen, 2012). De daaruit voortkomende database van bijwoordelijke bepalingen is vervolgens met behulp van Cesax (Komen, 2012) geannoteerd volgens de classificatie van Biber *et al.* (1999), waarin een onderscheid gemaakt wordt tussen verschillende semantische categorieën bijwoordelijke bepalingen die in drie hoofdcategorieën uiteenvallen: situerende bijwoordelijke bepalingen, bijwoordelijke bepalingen van modaliteit, en zinsverbindende bijwoordelijke bepalingen. Ook werd voor elke bijwoordelijke bepaling aangegeven of deze wel of niet referentieel werd gebruikt.

Hoofdstuk 2 onderzoekt bepalingen van tijd en toevoeging, belangrijke subcategorieën van respectievelijk situerende en zinsverbindende bijwoordelijke bepalingen. Beide worden in het Nederlands ook vaak referentieel gebruikt. De verwachting was dat Nederlandse leerders van het Engels dit in hun T2 ook meer zouden doen dan moedertaalsprekers. In vergelijking met Engelstalige schrijvers

van wetenschappelijke artikelen op het gebied van literatuur en taalkunde, gebruiken de Nederlandse leerders inderdaad meer bepalingen van plaats en toevoeging. En deze worden ook vaker referentieel gebruikt, door een verbinding te leggen met een antecedent in de direct voorafgaande zin.

Terwijl het aantal plaatsbepalingen tussen jaar 1 en 3 steeds verder afneemt en daarmee steeds meer op de frequentie in het T1 referentiecorpus gaat lijken, neemt het aantal bepalingen van toevoeging gestaag toe. Plaatsbepalingen komen in het Engels veel vaker aan het einde dan aan het begin van de zin voor en het lijkt erop dat de Nederlandse leerders zich daarvan tussen jaar 1 en 3 steeds meer bewust worden. Bepalingen van toevoeging, daarentegen, hebben logischerwijs een voorkeur voor het begin van de zin, zowel in het Nederlands als in het Engels, omdat ze een zinsverbindende functie hebben. Moedertaalsprekers van het Engels plaatsen deze bijwoordelijke bepaling dus ook aan het begin van de zin, maar zij gebruiken er minder. De toename van bepalingen van toevoeging zou daarom te maken kunnen hebben met het groeiende besef van de Nederlandse leerders dat het belangrijk is structuur aan te brengen in teksten, maar dat ze daarbij zwaar leunen op één enkele vertrouwde strategie, en andere mogelijkheden minder benutten dan de moedertaalsprekers. Er is wel een afname in het aantal bijwoordelijke bepalingen van toevoeging die een expliciet verbindende rol hebben. Toch is er in jaar 3 nog steeds een verschil tussen de leerders en de moedertaalsprekers, omdat de moedertaalsprekers hun bijwoordelijke bepalingen van toevoeging op een enkele uitzondering na nooit referentieel gebruiken.

Het onderwerp van hoofdstuk 3 is niet zozeer de longitudinale ontwikkeling van leerders van het Engels, maar de relatie tussen taalvaardigheidsniveau van Nederlandstalige eerstejaars studenten Engels en hun gebruik van initiële bijwoordelijke bepalingen. Het taalvaardigheidsniveau van de

leerders werd vooraf bepaald door middel van een *Oxford Online Placement Test* (OOPT). Een aanzienlijk deel van de eerstejaars werd door de OOPT ingeschat op C2, het hoogste niveau van het Europees Referentiekader (ERK). Deze uitslag geeft de indruk dat het ingangsniveau in jaar 1 gelijk is aan het uitstroomniveau aan het einde van jaar 3. Dit roept vanzelfsprekend vragen op over de toetsvaliditeit, maar het is hoe dan ook duidelijk dat ERK niveau C2 een heel groot gebied van taal en taalvaardigheid beslaat waarbinnen veel ruimte is voor variatie. Omdat de C-niveaus van het ERK alleen in grote lijnen zijn gedefinieerd, is het lastig om te bepalen waar precies binnen dat gebied een taalleerder zich bevindt. Het doel van dit hoofdstuk was dan ook om te onderzoeken of het mogelijk was de C-niveaus verder in te kleuren door de ontwikkeling van onze studenten te relateren aan hun ERK niveau.

Hiertoe werd eerst het aantal bijwoordelijke bepalingen in teksten geschreven door eerstejaars studenten vergeleken met teksten geschreven door moedertaalsprekers met vergelijkbare academische achtergrond. De statistische analyse wijst uit dat het gemiddelde aantal referentiële bijwoordelijke bepalingen significant hoger is in de teksten geschreven door de Nederlandse studenten dan in de teksten van de moedertaalsprekers. Er is ook een significante negatieve correlatie tussen het aantal referentiële bijwoordelijke bepalingen in een tekst en de OOPT score van de schrijver. Het gebruik van deze bijwoordelijke bepalingen zou dus opgenomen kunnen worden in een moeder- en doeltaalspecifieke aanvullende descriptor voor ERK niveau C2 met als doel vergevorderde taalontwikkeling nader te definiëren.

Hoofdstuk 4 presenteert een longitudinale statistische analyse van in totaal 780,414 woorden van syntactisch geannoteerde tekst en onderzoekt in hoeverre de frequentie waarmee initiële bijwoordelijke bepalingen gebruikt worden afhangt van moedertaal, genre, vakgebied, longitudinale ontwikkeling en

de tijdsdruk waaronder schrijfoopdrachten zijn uitgevoerd. Hieruit blijkt dat er in betogen meer zinsverbindende bijwoordelijke bepalingen voorkomen, terwijl in academische teksten over literaire onderwerpen meer situerende bijwoordelijke bepalingen gebruikt worden. Dit genreverschil gaat op voor zowel de T1 als de T2 teksten. In beide genres gebruikten T2 schrijvers over het geheel genomen meer bijwoordelijke bepalingen dan T1 schrijvers. Ze gebruikten ook meer bijwoordelijke bepalingen met een referentiële functie. Hiermee lijken ze hun Nederlandse manier om tekstuele samenhang te creëren toe te passen op het Engels. De gemiddelde frequentie waarmee de T2 schrijvers de verschillende types bijwoordelijke bepaling gebruiken komt in de loop van de jaren steeds dichterbij die van de T1 schrijvers te liggen, met uitzondering van de zinsverbindende bijwoordelijke bepalingen. Afhankelijk van het genre, blijft het gebruik daarvan gelijk, of neemt het zelfs toe. Verder lijkt tijdsdruk bepalend voor de mate waarin de T2 teksten de T1 teksten benaderen. Bij teksten die niet onder tijdsdruk geschreven waren, lag de frequentie van alle categorieën bijwoordelijke bepalingen dichterbij die van de T1 teksten. Dit doet vermoeden dat studenten onder tijdsdruk terugvallen op strategieën om tekststructuur aan te brengen waarmee ze vanuit hun moedertaal bekend zijn.

Hoofdstuk vijf onderzoekt twee mogelijke verklaringen voor de sterke neiging van Nederlandse leiders van het Engels om teksten te structureren door middel van initiële bijwoordelijke bepalingen: moedertaalinterferentie en onderwijs. Een contrastieve analyse toonde aan dat Nederlandse studenten Engels in hun Nederlands nog meer initiële bijwoordelijke bepalingen gebruiken dan in hun Engels. Dit ondersteunt de veronderstelling dat interferentie vanuit het Nederlands een rol speelt in het gebruik van bijwoordelijke bepalingen van Nederlandse leiders. Het verklaart echter niet waarom de invloed van de moedertaal wel zichtbaar is in het veelvuldig gebruik van zinsverbindende en

referentiële bijwoordelijke bepalingen in de T2 teksten, terwijl andere in het Nederlands veelvoorkomende bijwoordelijke bepalingen, m.n. die van modaliteit, niet doorsijpelen in het Engels.

Onderwijs zou ook een rol kunnen spelen in de frequentie waarmee Nederlandse leeders bijwoordelijke bepalingen gebruiken. Om dit nader te onderzoeken, bestudeerden we hoe zinsverbindende bijwoordelijke bepalingen in veelgebruikte lesmethodes Engels voor het VWO worden behandeld. Vaak worden deze gepresenteerd in lijstjes zonder context, gegroepeerd per semantische categorie. Daardoor zou leeders de indruk kunnen krijgen dat bijwoordelijke bepalingen die tot dezelfde categorie behoren onderling uitwisselbaar zijn. Bovendien worden leeders zeer regelmatig aangespoord zinsverbindende bijwoordelijke bepalingen te gebruiken om hun teksten te structureren. Andere strategieën om tekstuele samenhang te creëren komen weinig aan bod. Het herkennen van zinsverbindende bijwoordelijke bepalingen, veelal aangeduid als ‘signaalwoorden’, wordt ook gezien als een belangrijke leesstrategie voor het centraal schriftelijk eindexamen. In examentraining zowel op scholen als daarbuiten wordt hieraan veel aandacht besteed. Hoewel er dus inderdaad overeenkomst is tussen het veelvuldig gebruik van zinsverbindende bijwoordelijke bepalingen door studenten Engels in hun eerste en hun tweede taal, wat zou kunnen wijzen op moedertaalinterferentie, wordt dit effect mogelijk versterkt door een enigszins reductionistische benadering van tekststructuur en cohesie in het onderwijs.

Als interferentie werkelijk ten grondslag ligt aan het gebruik van bijwoordelijke bepalingen door Nederlandse leeders, is de verwachting dat er een duidelijk verschil zichtbaar moet zijn tussen het Engels dat zij produceren en dat van leeders met een andere moedertaalachtergrond. Dit is de achterliggende gedachte van de analyse van initiële bijwoordelijke bepalingen in het Engels van

Nederlandstalige en Franstalige studenten Engels in hoofdstuk 6. De Nederlandstalige studenten Engels bleken bepaald niet uniek te zijn in de frequentie waarmee zij initiële bijwoordelijke bepalingen gebruiken. De Franstalige studenten gebruiken er zelfs nog meer. Op basis van de kwantitatieve analyse lijkt het veelvuldig gebruik van initiële bijwoordelijke bepalingen dus niet het gevolg te zijn van interferentie vanuit het Nederlands.

De kwalitatieve analyse, daarentegen, maakte duidelijk dat de Nederlandstalige en Franstalige leerders weliswaar op elkaar lijken in de frequentie waarmee ze initiële bijwoordelijke bepalingen gebruiken, maar van elkaar verschillen in de manier waarop ze die realiseren. Terwijl referentieel gebruik van initiële bijwoordelijke bepalingen kenmerkend is voor de Nederlandse leerders, lijken de Franstalige leerders vaker bijwoordelijke bepalingen in de eerste positie te plaatsen uit stilistische overwegingen, bijvoorbeeld om variatie aan te brengen in zinsstructuur, om een ritmisch effect teweeg te brengen door middel van een reeks bijwoordelijke bepalingen, of om balans in de zin aan te brengen door het gewicht van een bijwoordelijke bepaling naar het begin van de zin te halen. Moedertaalinterferentie lijkt dus toch een rol te spelen in de manier waarop initiële bijwoordelijke bepalingen gebruikt worden en de pragmatische functies die ze vervullen.

De conclusie lijkt gerechtvaardigd dat het geschreven Engels van vergevorderde Nederlandstalige leerders gekenmerkt wordt door de frequentie waarmee ze initiële bijwoordelijke bepalingen gebruiken die tot doel hebben de interne samenhang van de tekst te vergroten, namelijk zinsverbindende bijwoordelijke bepalingen en bijwoordelijke bepalingen – van alle semantische categorieën – die referentieel gebruikt worden. Hier lijken twee verschillende oorzaken voor te zijn. Het is aannemelijk dat interferentie op het snijvlak van syntaxis en pragmatiek ten grondslag ligt aan het gebruik van referentiële

bijwoordelijke bepalingen waarmee zinnen in de voorafgaande tekst verankerd worden. Daarvoor zijn drie soorten bewijs: a) intra-T1-groep homogeniteit, b) inter-T1-groep heterogeniteit en c) T1-T2 congruentie (Jarvis, 2000). Intra-T1-groep homogeniteit is zichtbaar in het feit dat verschillende cohorten Nederlandstalige studenten Engels consequent bijwoordelijke bepalingen gebruiken om de zin in de voorafgaande context te verankeren. Dit is onafhankelijk van het teksttype. Er is inter-T1-groep heterogeniteit in het gebruik van referentiële bijwoordelijke bepalingen tussen vergevorderde Nederlandstalige leerders aan de ene kant en zowel moedertaalsprekers als Franstalige leerders aan de andere. Er is ook sprake van T1-T2 congruentie, in die zin dat het gebruik van referentiële bijwoordelijke bepalingen van Nederlandse leerders van het Engels niet alleen zichtbaar is in hun tweede taal, maar ook in hun moedertaal.

Maar interferentie biedt geen bevredigende verklaring voor het veelvuldig gebruik van zinsverbindende bijwoordelijke bepalingen door Nederlandstalige leerders, ondanks de T1-T2 congruentie die ook voor deze categorie bijwoordelijke bepalingen bestaat tussen T1 Nederlandse en T2 Engelse teksten. Ten eerste doet de forse toename in het aantal zinsverbindende bijwoordelijke bepalingen in essays over literaire onderwerpen vermoeden dat het gebruik hiervan niet zozeer het gevolg is van onbewuste moedertaalinterferentie, maar eerder van een bewuste inspanning om tekststructuur aan te brengen. Een ander bewijs dat tegen de rol van interferentie spreekt is het feit dat niet alleen de Nederlandstalige maar ook de Franstalige leerders meer zinsverbindende bijwoordelijke bepalingen gebruiken dan de moedertaalsprekers. Het idee dat taalleerders met uiteenlopende moedertaalachtergronden opvallend veel zinsverbindende bijwoordelijke bepalingen gebruiken komt overeen met eerdere studies, zoals Crewe (1990), Field & Yip (1992), Milton & Tsang (1993) en Leńko-Szymańska (2008). In elk van deze onderzoeken wordt de nadruk die in het

schrijfonderwijs op dit soort verbindingswoorden wordt gelegd als mogelijke verklaring genoemd voor de misvatting van taalleerders dat er een directe positieve correlatie bestaat tussen het aantal verbindingswoorden dat ze gebruiken en de coherentie van hun tekst.

In de Nederlandse context wordt deze verklaring bevestigd door de reductionistische benadering van tekstuele cohesie in veelgebruikte lesmethodes Engels, waarin de nadruk op verbindingswoorden ten koste lijkt te gaan van een representatieve behandeling van andere strategieën om tekstuele samenhang te creëren. In het schoolvak Nederlands ligt ook veel nadruk op logische relaties en de signaalwoorden waarmee deze worden aangeduid. Als het gebruik van initiële zinsverbindende bijwoordelijke bepalingen in het Engels van Nederlandstalige leerders een weerspiegeling is van het gebruik hiervan in hun T1, Nederlands, is dit dus niet noodzakelijkerwijs, of in ieder geval niet exclusief, het gevolg van onbewuste moedertaalinterferentie, maar eerder van bewuste toepassing van aangeleerde schrijfconventies.

Curriculum Vitae

Sanne van Vuuren (Amsterdam, 1981) studied English language and culture at Radboud University in Nijmegen, the Netherlands, graduating with distinction in 2003. In the years that followed, she combined work as a language trainer at the university's language and communication centre, Radboud in'to Languages, with teaching oral communication skills, writing, grammar and translation courses at the department of English language and culture. Her teaching inspired an interest in second language acquisition, eventually motivating her to embark on a PhD at the Nijmegen Centre for Language Studies (CLS) in 2012. Sanne continues to work at the department of English Language and Culture, combining her interest in second language acquisition research and her love of teaching as a lecturer in English Language and Linguistics.