

Universität Siegen
Philosophische Fakultät

**Dissertation zur
Erlangung des Doktorgrades
im Fach Englische Sprachwissenschaft**

Prominence assignment in English triconstituent compounds

Kristina Kösling

Gutachter:
Prof. Dr. Ingo Plag
Prof. Dr. Heinz J. Giegerich

Eingereicht am 09.11.2011
Tag der mündlichen Prüfung: 28.03.2012

Gedruckt auf alterungsbeständigem holz- und säurefreiem Papier.

Contents

1	Introduction	1
2	Prominence assignment in English nominal compounds	7
2.1	Prominence assignment in biconstituent Noun+Noun compounds	7
2.2	Prominence assignment in triconstituent Noun+Noun+Noun compounds	14
3	General Methodology	23
3.1	How to determine prominence patterns of English compounds?	23
3.2	Pitch measurements	27
3.3	Discourse factors	32
3.4	Mixed-effects models	33
4	Corpus Study	37
4.1	Methodology	37
4.1.1	The data: The Boston University Radio Speech Corpus	37
4.1.2	Branching direction	40
4.1.3	Acoustic measurements	42
4.1.4	Prediction: Lexical Category Prominence Rule	43
4.1.5	Statistical procedure	46
4.2	Results	51
4.2.1	Testing the LCPR	51
4.2.2	Variation and branching direction	55
4.3	Discussion	61
4.3.1	Violations of the LCPR at the N-level	61
4.3.2	Violations of the LCPR at the IC-level	63
4.4	Summary and conclusion	66
5	Experimental data	69
5.1	Introduction	69
5.2	Common methodology: Experiments 1 and 2	71

5.2.1	Subjects	71
5.2.2	Experimental design	72
5.2.3	Recording procedure	74
5.2.4	Acoustic measurements	74
6	Production Experiment 1: Embedded Prominence Hypothesis (EPH)	77
6.1	Methodology	77
6.1.1	Stimuli	77
6.1.2	Embedded Prominence Hypothesis: predictions	81
6.1.3	Statistical procedure	85
6.2	Results	91
6.2.1	EPH: Embedded left prominent NN compounds	92
6.2.2	EPH: Embedded right prominent NN compounds	95
6.3	Discussion	100
6.4	Individual speaker variation	104
6.5	Conclusion	108
7	Production Experiment 2: IC-Prominence Hypothesis (IPH)	111
7.1	Methodology	111
7.1.1	Stimuli	111
7.1.2	Statistical procedure	116
7.2	Results	118
7.2.1	Analogy: <i>avenue</i> vs. <i>street</i>	118
7.2.2	Analogy: <i>pie</i> vs. argument-head compounds	121
7.2.3	Semantics: <i>IC2 during IC1</i>	125
7.2.4	Semantics: <i>IC2 is located at IC1</i>	129
7.2.5	Individual item variation	131
7.3	Discussion	141
7.3.1	Analogy	141
7.3.2	Semantics	144
7.4	Conclusion	146
8	Summary and conclusion	149
A	Data	155
B	Mixed-effects models, regression models and likelihood ratio tests	175

List of Tables

2.1	Semantic subgroups which trigger right prominence in NN compounds	9
4.1	Corpus data: number of types sampled across speakers	40
4.2	Mixed-effects model: left-branching vs. right-branching compounds. Baseline: left, mpitch, female	52
4.3	Violations of the LCPR: left-branching compounds	58
4.4	Violations of the LCPR: right-branching compounds	59
4.5	Semantic relations at the IC-level of some right-branching compounds with highest prominence on IC2	65
6.1	Types of compounds constructed to test the EPH	78
6.2	Example of each compound type constructed to test the EPH	80
6.3	Mixed-effects model: L/N1 vs. L/N2 vs. R/N2 vs. R/N3 compounds. Baseline: ln1, mpitch, female	94
6.4	Mixed-effects model: L/N1 vs. L/N2 vs. R/N2 vs. R/N3 compounds. Baseline: lrn2, rpitch, female	97
6.5	Mixed-effects model: L/N1 vs. L/N2 vs. R/N2 vs. R/N3 compounds. Baseline: lrn2, lpitch, female	100
6.6	Excluded speaker: Mean values and Standard deviations of pitch values for L/N1, L/N2, R/N2 and R/N3 compounds.	106
7.1	List of <i>avenue</i> and <i>street</i> compounds	113
7.2	List of <i>pie</i> and argument-head compounds	114
7.3	List of <i>IC2 during IC1</i> and argument-head compounds	115
7.4	List of <i>IC2 is located at IC1</i> and argument-head compounds	115
7.5	Mixed-effects model: <i>avenue</i> vs. <i>street</i> compounds. Baseline: <i>avenue</i> , lpitch, female	119
7.6	Mixed-effects model: <i>avenue</i> vs. <i>street</i> compounds. Baseline: <i>avenue</i> , mpitch, female	119

List of Tables

7.7	Mixed-effects model: <i>avenue</i> vs. <i>street</i> compounds. Baseline: <i>avenue</i> , <i>rpitch</i> , <i>female</i>	119
7.8	Mixed-effects model: <i>pie</i> vs. <i>argument-head</i> compounds. Baseline: <i>pie</i> , <i>lpitch</i> , <i>female</i>	123
7.9	Mixed-effects model: <i>pie</i> vs. <i>argument-head</i> compounds. Baseline: <i>pie</i> , <i>mpitch</i> , <i>female</i>	123
7.10	Mixed-effects model: <i>pie</i> vs. <i>argument-head</i> compounds. Baseline: <i>pie</i> , <i>rpitch</i> , <i>female</i>	123
7.11	Mixed-effects model: <i>IC2 during IC1</i> compounds vs. <i>control</i> items. Baseline: <i>duration</i> , <i>mpitch</i> , <i>female</i>	127
7.12	Mixed-effects model: <i>IC2 is located at IC1</i> compounds vs. <i>control</i> items. Baseline: <i>location</i> , <i>mpitch</i> , <i>female</i>	130

List of Figures

2.1	Metrical tree: left- and right-branching compound	15
4.1	Pitch track of a left-branching compound: <i>[crédit card] companies</i>	45
4.2	Pitch track of a right-branching compound: <i>corner [drúg store]</i>	45
4.3	Quantile-quantile plots for pitch distribution grouped by speakers	48
4.4	Interaction plot of mixed-effects model: BURSC	52
4.5	Pitch track and metrical tree: <i>[science fiction] shocker</i>	58
4.6	Pitch track and metrical tree: <i>[child care] crisis</i>	59
4.7	Pitch track and metrical tree: <i>state [health programs]</i>	60
4.8	Pitch track and metrical tree: <i>China [information center]</i>	60
6.1	Pitch track of L/N1 compound: <i>[háy fever] treatment</i>	82
6.2	Pitch track of R/N2 compound: <i>piano [shéet music]</i>	82
6.3	Pitch track of L/N2 compound: <i>[science fiction] shocker</i>	84
6.4	Pitch track of R/N3 compound: <i>student [string órchestra]</i>	85
6.5	Standardized residuals against fitted values, full model	88
6.6	Standardized residuals against fitted values, trimmed model	88
6.7	Interaction plot: L/N1, L/N2, R/N2, R/N3 compounds	91
6.8	Interaction plot: L/N1 vs. R/N2 compounds, female speakers	93
6.9	Interaction plot: R/N3 vs. R/N2 compounds, female speakers	97
6.10	Interaction plot excluded speaker: L/N1, L/N2, R/N2, R/N3	105
7.1	Interaction plot: <i>avenue</i> vs. <i>street</i> compounds	120
7.2	Interaction plot: <i>pie</i> vs. argument-head compounds	124
7.3	Interaction plot: <i>IC2 during IC1</i> compounds vs. control items	127
7.4	Interaction plot: <i>IC2 is located at IC1</i> compounds vs. control items	130
7.5	Quantile-quantile plot: residuals of the <i>IC2 during IC1</i> model	131
7.6	Quantile-quantile plot: residuals of the <i>IC2 is located at IC1</i> model	131
7.7	Interaction plot: <i>holiday season job</i> vs. <i>holiday season planner</i>	134
7.8	Interaction plot: <i>winter term breakfast</i> vs. <i>winter term preparation</i>	134

List of Figures

7.9	Interaction plot: <i>flight test accident</i> vs. <i>flight test manual</i>	134
7.10	Interaction plot: <i>field trip weather</i> vs. <i>field trip leader</i>	134
7.11	Interaction plot: <i>monday morning meeting</i> vs. <i>monday morning hater</i>	134
7.12	Interaction plot: <i>spring break vacation</i> vs. <i>spring break organization</i>	134
7.13	Interaction plot: <i>cocktail bar fight</i> vs. <i>cocktail bar designer</i>	138
7.14	Interaction plot: <i>movie theatre fire</i> vs. <i>movie theatre visitor</i>	138
7.15	Interaction plot: <i>gas station robbery</i> vs. <i>gas station owner</i>	138
7.16	Interaction plot: <i>coffe house concert</i> vs. <i>coffee house lover</i>	138
7.17	Interaction plot: <i>apartment building party</i> vs. <i>apartment building manager</i>	138
7.18	Interaction plot: <i>hockey stadium event</i> vs. <i>hockey stadium cleaner</i>	138

1 Introduction

It has been traditionally argued that compounds in English are generally left prominent (e.g. Chomsky and Halle 1968). Later research (e.g. Fudge 1984; Liberman and Sproat 1992; Olsen 2000; Olsen 2001; Plag et al. 2008) has shown, however, that prominence assignment in compounds is quite variable, at least with regard to NN compounds, the group of compounds which has been the primary focus of that research. In contrast, less attention has been paid to the prominence behaviour of larger compounds and in particular the group of triconstituent Noun+Noun+Noun compounds (NNN). This fact is fairly surprising given that the prominence pattern of NNN compounds has been claimed to be basically governed by the same phonological rule that operates in NN compounds and which often fails. In view of this situation, the current thesis takes a closer look at the generalizations regarding prominence assignment in NNN compounds.

Triconstituent NNN compounds can be subdivided into three different types according to their internal structure, namely left-branching compounds (e.g. *[seat belt] law*), right-branching compounds (e.g. *team [locker room]*) and structurally ambiguous compounds (e.g. *[kitchen towel] rack* or *kitchen [towel rack]*). It is generally assumed that the prominence pattern of triconstituent compounds depends on the compounds' internal structure, i.e. the branching direction of the compounds; in left-branching compounds highest prominence is assigned to the leftmost constituent, whereas in right-branching compounds the second constituent of the entire compound is the most prominent one. For structurally ambiguous compounds it is assumed that highest prominence is either assigned to constituent N1 or constituent N2, depending on whether the compound is semantically interpreted as left- or right-branching by the speaker. The generalizations regarding prominence assignment in triconstituent compounds are for instance captured in Liberman and Prince's (1977) 'Lexical Category Prominence Rule' (LCPR) and illustrated with the examples given in (1) and (2) below.

(1) **Left-branching compounds**

[séat belt] law

[dáy care] space

[láu degree] requirement

[béach ball] game

(2) **Right-branching compounds**

team [lócker room]

morning [néwspaper]

Boston [gáing members]

beach [báall game]

However, a closer look at the relevant literature reveals that the generalizations regarding prominence assignment to NNN compounds are not the result of large-scale empirical investigations. Instead, the assumptions seem to be based on the analysis of small sets of self-selected, mostly isolated examples. These examples are frequently repeated throughout the literature to illustrate the typical prominence pattern predicted for left- and right-branching compounds. Furthermore, thus far the assignment of prominence to triconstituent compounds has been primarily based on the researchers' own intuition about prominence (e.g. Kingdon 1958; Liberman and Prince 1977; Giegerich 1985; Giegerich 1992; Carstairs-McCarthy 2002; Berg 2009). However, this approach to determine prominence in compounds has been argued to be problematic given that speakers may vary in the assignment of prominence to one and the same compound (e.g. Bauer 1983a; Plag 2006; Kunter 2011).

Apart from that, the literature also provides some counter-examples to the rule (e.g. Kingdon 1958; Kvam 1990; Berg 2009; Giegerich 2009); left- and right-branching compounds have been cited with highest prominence on the first, second and third constituent, irrespective of the compounds' branching direction. Some examples are given in (3) for a better illustration.¹

(3) **Counter-examples to the LCPR**

ówl [nest box]

university [spring térm]

[living-room] fúrniture

[garden shéd] exhibition

Although the counter-examples to the LCPR, as found in the literature, also primarily rest on individual speaker judgements, the existence of such violations does not seem surprising if one considers the underlying assumptions on which the LCPR, i.e. the factor branching direction, rests. The LCPR is based on the assumption that

¹The examples given in (3) are taken from Giegerich (2009:10).

Noun+Noun (NN) compounds in English are categorically left prominent (e.g. *séat belt* in *séat belt law*; *lócker room* in *team lócker room*). However, recent experimental and corpus studies by Plag (2006), Plag et al. (2007), Lappe and Plag (2007), Plag et al. (2008), Bell (2008), Plag and Kunter (2010) and Kunter (2011) dealing with prominence assignment to biconstituent compounds have shown that a considerable number of right prominent NN compounds exist in English as well. Given this variable prominence behaviour of NN compounds, which function as complex constituents in triconstituent compounds, there is ample reason to assume that right prominent NN compounds may also affect prominence assignment in triconstituent NNN compounds.

Yet, surprisingly to my knowledge, except for a very recent theoretically based approach by Giegerich (2009), the potential influence of right prominent NN compounds on prominence assignment in triconstituent NNN compounds has never been explicitly discussed in the literature nor has it been empirically investigated with large amounts of independently gathered data. Instead, the potential consequences that the existence of right prominent NNs may have for prominence assignment in larger compounds have been largely ignored in the relevant literature.

Hence, given this lack of empirical support for the LCPR as well as the problematic assumptions on which the generalizations are based, the central aim of this thesis is to determine how triconstituent NNN compounds are really stressed in English, with a focus on left- and right-branching compounds. The thesis presents a first systematic investigation of the prominence behaviour of English triconstituent NNN compounds analyzing both large amounts of speech corpus data and experimentally obtained data. On the one hand, it addresses the question to what extent the LCPR actually holds for left- and right-branching compounds. On the other hand, it presents the first study to empirically investigate whether the existence of right prominent NN compounds affects prominence assignment in triconstituent NNN compounds. In addition to that, the present thesis differs from previous accounts dealing with prominence assignment in triconstituent compounds with reference to the methodology applied to account for the prominence behaviour of triconstituent NNN compounds. As mentioned above, thus far prominence assignment to triconstituent compounds has been based on the researchers' own intuition about stress and that of a few additional judges. In the present thesis this rather subjective way of determining prominence in NNN compounds is replaced by a more objective approach, namely that of measuring pitch as an acoustic cue to prominence.² Such a phonetically based approach

²When measuring pitch, one measures the fundamental frequency (F0) of the speech signal. The F0 is

to compound stress has already been successfully employed in experimental and corpus studies dealing with prominence assignment in NN compounds (e.g. Plag 2006; Plag et al. 2007; Plag et al. 2008; Kunter 2011) and can be shown to be also appropriate in order to determine prominence differences between left- and right-branching triconstituent NNN compounds.

The thesis is structured as follows. Chapter 2 provides the theoretical background for the study. The chapter starts with a review of the current state of research regarding prominence assignment in biconstituent NN compounds. This review of the factors responsible for prominence assignment in NN compounds is essential for an understanding of the problems associated with the generalizations regarding prominence assignment in NNN compounds. The second part of chapter 2 focuses on the generalizations regarding prominence assignment in triconstituent NNN compounds. It presents a detailed discussion of the theoretical shortcomings associated with Liberman and Prince's (1977) Lexical Category Prominence Rule and discusses findings and shortcomings of previous studies dealing with the prosodic structure of NNN constructions. The discussion ends with the formulation of a research question and two new hypotheses regarding prominence assignment in triconstituent compounds, i.e. the Embedded Prominence Hypothesis (EPH) and the IC-Prominence Hypothesis (IPH). The two hypotheses differ from the LCPR in that they incorporate the existence of right prominent NN compounds into their predictions. The research question is addressed in a corpus study, whereas the two hypotheses are tested by means of two production experiments.

Chapter 3 describes the general methodology by means of which the prominence patterns of the triconstituent compounds are determined in both the corpus and the experimental studies. As noted earlier in this section, in contrast to previous studies, I decided to use automatic measurements of pitch in order to determine prominence in triconstituent compounds. Chapter 3 introduces this method and discusses various problems associated with measuring the fundamental frequency (F0), i.e. the acoustic correlate of pitch. The chapter closes with a section on mixed-effects models, which provide the statistical tool by means of which the prominence differences between the different types of compounds are statistically modelled in this thesis.

Chapter 4 presents a corpus study which addresses the question to what extent the predictions of the LCPR actually hold for left- and right-branching compounds in English. The corpus study described in chapter 4 is a more elaborate version of the corpus study presented in Kösling and Plag (2009). Given these circumstances, it

the acoustic correlate of what listeners perceive as pitch (e.g. Ladefoged 2003).

should be noted that several parts of chapter 4 have already been published as such or in a slightly different version in Kösling and Plag (2009). This particularly refers to subsections 4.1.1, 4.1.2, 4.2.2 and section 4.3.³

The thesis continues with chapter 5, which provides some general information about the two production experiments designed in order to test the Embedded Prominence Hypothesis and the IC-Prominence Hypothesis. In particular, the chapter includes details on the speakers participating in the two experiments, the recording procedure, the general test design of the experiments and the annotation process of the data. Chapter 6 describes the first production experiment, which tested the predictions of the Embedded Prominence Hypothesis. The chapter provides information on the stimuli constructed in order to test the EPH and discusses some further methodological issues specific to the experiment. Furthermore, it presents the result of the experiment and a thorough discussion of it before ending with some first concluding thoughts. Chapter 7 describes the second production experiment specifically designed in order to test the IC-Prominence Hypothesis. The chapter introduces the stimuli constructed in order to test the IPH and describes the statistical modelling of the data. Furthermore, it presents the result of the analysis, its discussion and draws some first conclusions. The thesis ends with a summary and a general conclusion in chapter 8.

³The same holds true for large parts of section 2.2 and some paragraphs of chapter 3 of this thesis.

2 Prominence assignment in English nominal compounds

2.1 Prominence assignment in biconstituent Noun+Noun compounds

The majority of English Noun+Noun compounds seems to be stressed as predicted by Chomsky and Halle's (1968) 'Compound Stress Rule' (CSR). The CSR predicts that English Noun+Noun compounds are left prominent as illustrated with the examples in (1).¹

- (1) *páyment problem*
installátion guide
spáce requirement

However, while the majority of NN compounds behaves as predicted by the CSR, the literature has also provided a considerable number of counter-examples to the rule, i.e. Noun+Noun compounds with highest prominence on the right-hand constituent (e.g. Kingdon 1958; Fudge 1984; Ladd 1984; Liberman and Sproat 1992; Olsen 2000, 2001; Plag 2006; Plag et al. 2007, 2008; Lappe and Plag 2007; Plag and Kunter 2010; Plag 2010; Kunter 2011). A few right prominent Noun+Noun compounds are given in (2) for a better illustration.

- (2) *Madison Ávenue*
Boston márathom
geologist-astrónomer

¹The examples given in (1) and (2) are taken from Plag (2003:137f).

The examples in (2) have raised the question of which factors are responsible for the aberrant prominence behaviour to the CSR and various different approaches have been proposed in the literature trying to explain this variation. One such approach is that right prominent NNs are not compounds but syntactic phrases (e.g. Chomsky and Halle 1968; Marchand 1969). One reason for this assumption is that syntactic phrases, in contrast to compounds, are assumed to be generally right prominent in English, with their prominence pattern governed by Chomsky and Halle's (1968) 'Nuclear Stress Rule' (NSR). However, this approach of explaining prominence variation in NN compounds has been seriously questioned by various other linguists (e.g. Bauer 1998; Olsen 2000; Bell and Plag 2012) who point to a lack of evidence other than the compounds' prominence pattern that would support a phrasal analysis of right prominent NNs. Because of this absence of convincing criteria to differentiate between NN compounds and NN phrases many linguists have favoured an analysis in which both left and right prominent NNs are treated as compounds (e.g. Fudge 1984; Ladd 1984; Olsen 2000; Bauer 1998; Giegerich 2004; Bell and Plag 2012). According to these authors the variable prominence behaviour of Noun+Noun compounds is triggered by either semantic factors, structural differences, analogical mechanisms or informativeness (cf. e.g. Plag and Kunter 2010). In various experimental and corpus studies by Plag (2006), Plag et al. (2007), Plag et al. (2008), Lappe and Plag (2007), Plag and Kunter (2010), Plag (2010), Kunter (2011) and Bell and Plag (2012), using regression and logistic analysis as well as exemplar based models, it turned out that of these different factors only analogy, semantics and informativeness are significant predictors, with the factor structure being less capable of accounting for the variable prominence behaviour of NN compounds.²

In particular, a group of linguists (e.g. Fudge 1984; Ladd 1984; Sproat 1994; Olsen 2000; Olsen 2001) has argued that right prominence is dependent on the semantic properties of the compounds' left and right constituent or the semantic relation between the two compound constituents.³ For instance, it has been argued that compounds in which *N1 is a proper noun* (e.g. *Madison Park*) or *N2 is a geographical term*

²The structural approach (e.g. Giegerich 2004) argues that prominence assignment to compounds is dependent on the structural make-up of the compound, with modifier-head compounds being right prominent and complement-head compounds being generally left prominent. Yet, according to studies by Plag and his colleagues the effect of 'structure' is restricted to complement-head compounds with a deverbal head ending in *-er*. These compounds showed a statistical tendency towards left prominence. For more details on this approach, the reader is referred to e.g. Plag (2006), Plag et al. (2008) or Kunter (2011).

³This approach of explaining prominence variation among NN compounds has been referred to as the 'semantic hypothesis' in studies by Plag and his colleagues. For a more thorough discussion, the reader is referred to e.g. Plag (2006), Plag et al. (2007) or Plag et al. (2008).

(e.g. *Sun Valley*) are categorically right prominent. The same prominence behaviour has also been associated with compounds which can be semantically interpreted as *N2 is located at N1* (e.g. *Boston harbour*), *N2 is during N1* (e.g. *summer job*), *N2 is made of N1* (e.g. *silk dress*) or *N2 is N1* (e.g. *geologist-astronomer*). The semantic approach was tested in studies by Plag (2006), Plag et al. (2007, 2008), Lappe and Plag (2007), Plag and Kunter (2010), Plag (2010), Kunter (2011) and Bell and Plag (2012). Whereas the experimental study by Plag (2006) revealed conflicting results for the semantic factor, the remaining corpus-based studies showed effects in the predicted direction for some, yet not all, semantic relations and categories under investigation.⁴ Table 2.1 lists some of these semantic relations and categories which emerged as significant in the above mentioned studies. Yet, regarding the effects obtained for the semantic subgroups listed in Table 2.1 it must also be noted that they were gradient rather than categorical in nature. Thus, compounds belonging to one of these semantic subgroups were not categorically right prominent but rather showed a greater tendency towards right prominence than compounds not belonging to one of these semantic subcategories.

Table 2.1: Semantic subgroups which trigger right prominence in NN compounds

Semantic relations and categories	Studies investigating the semantic factor
N2 is made of N1	Plag et al. (2007, 2008); Plag (2010); Kunter (2011); Bell and Plag (2012)
N2 is during N1	Plag et al. (2007); Kunter (2011); Bell and Plag (2012)
N2 is located at N1	Plag et al. (2008); Kunter (2011); Bell and Plag (2012)
N1 is N2	Plag et al. (2007, 2008); Kunter (2011)
N1 has N2	Plag et al. (2008); Plag (2010); Kunter (2011)
N1 is a proper noun	Plag et al. (2007, 2008); Kunter (2011)

A second group of authors has argued that analogical mechanisms are the governing force behind the variable prominence behaviour in NN compounds (e.g. Schmerling 1971; Liberman and Sproat 1992; Spencer 2003). Proponents of this approach assume that compounds with the same left or right constituent also tend to exhibit the same prominence pattern. Examples that are usually mentioned in this context are that of *street vs. avenue* and *pie vs. cake* compounds (e.g. Ladd 1984; Liberman and Sproat 1992; Plag 2003). In particular, compounds with *cake* and *street* as their right-hand constituent are claimed to be generally left prominent (e.g. *apple cake, cherry cake,*

⁴Except for the two studies by Plag (2006) and Bell and Plag (2012), the different studies mentioned above tested the same set of semantic relations and categories. For an overview of these semantic subcategories, the reader is referred to Plag et al. (2007, 2008).

Óxford Street, Máin Street), whereas corresponding compounds with *avenue* and *pie* as head constituent are generally right prominent (e.g. *apple pie, cherry pie, Fifth Avenue, Madison Avenue*). In these examples, it is assumed that prominence is assigned in analogy to the head constituent.

The effect of analogy was tested, alongside other predictors, in studies by Plag (2006), Plag et al. (2007), Lappe and Plag (2007) and Plag and Kunter (2010). In Plag et al. (2007) the analogical influence of the left and right constituent was tested with compounds extracted from the lexical data base CELEX (Baayen et al. 1995) and the computational learning algorithm TiMBL (Daelemans et al. 2004). Plag et al. (2007) coded their data for semantic and structural features, the constituent family of the left and right constituent⁵ and the compounds' prominence pattern as provided by CELEX. The features were used by TiMBL to predict the prominence patterns of the compounds extracted from the CELEX data base.⁶ Plag et al. (2007) found that when TiMBL had no access to the information about the constituent family of the compounds, the predictive accuracy of the model dropped significantly (cf. Plag et al. 2007:225). In contrast, no such decrease in the predictive accuracy of the model was observed when information about semantics or structure of the compounds were excluded from the model, which indicated that the constituent family provides a strong predictor of prominence assignment in NN compounds.

A strong effect for the constituent family, i.e. the influence of analogy, was also determined in the study by Lappe and Plag (2007), which used two different computational algorithms, i.e. TiMBL and AM. Lappe and Plag (2007) found that analogical models which only used information about the constituent family had a higher predictive accuracy than models only working on the basis of semantic and structural features.

Finally, in a more recent study by Plag (2010) the potential influence of the constituent family on prominence assignment in NN compounds was investigated by analysing data from three different corpora, i.e. the lexical data base CELEX (Baayen

⁵The constituent family is the set of compounds that share the same left or right constituent with each other. For example, the compounds *summer job, summer dress* and *summer game* belong to the same left constituent family, whereas compounds like *winter term, summer term, spring term* are members of the same right constituent family (cf. e.g. Bell 2008).

⁶TiMBL assigns prominence to new compounds based on similarities to other compounds, which have been previously stored in the model's memory. In particular, for every new input form, TiMBL searches at first its memory for forms with similar features as the respective input form. These forms are the so-called nearest neighbours to the input. In a second step the algorithm assigns the input form the same prominence pattern as that exhibited by the majority of its nearest neighbours (cf. Plag et al. 2007:222) For more details on how TiMBL works, the reader is referred to Daelemans et al. (2004).

et al. 1995), the Boston University Radio Speech Corpus (Ostendorf et al. 1996) and a third corpus compiled by Teschner and Whiteley (2004). On the one hand Plag (2010) tested as to whether compounds that shared the same left or right constituent also exhibited the same prominence pattern. On the other hand, he also addressed the question whether the semantic effects found in Plag et al. (2007) and Plag et al. (2008) are independent effects or should be regarded as an epiphenomenon of the constituent family bias (cf. Plag 2010:248), which provides a measure for the tendency of a given constituent family to favour a particular prominence pattern (cf. Plag 2010:222).

For each of the three corpora Plag (2010) devised three different regression models; one model with the constituent family bias as the only predictor, one model with only semantic and structural predictors and a third model with constituent family bias and the structural and semantic predictors. According to Plag's analysis, the models with constituent family bias as the only predictor already revealed robust effects for this factor. More importantly, however, the constituent family bias also remained a robust and significant effect in those models which included all other predictors. In these models also some of the semantic relations and categories that had already been found to be significant in the studies by Plag et al. (2007) and Plag et al. (2008) emerged as independent effects in the analysis, yet with a weaker effect size as the constituent family bias. Given this result, Plag concluded that both constituent family and semantics operate alongside each other, yet with the constituent family bias being a stronger predictor than semantics (cf. Plag 2010:271f).

The third factor that has only recently been empirically investigated and found to affect prominence assignment in NN compounds is that of informativeness. The two corpus-based studies by Plag and Kunter (2010) and Bell and Plag (2012) investigated the assumption whether informative constituents are more likely to attract prominence than uninformative constituents. Whereas the study by Plag and Kunter (2010) revealed only weak effects for this factor, Bell and Plag (2012) provided strong effects for informativeness as a determinant of compound stress. As noted by Bell and Plag (2012:37) one reason for the discrepancy in the results of the two studies is that in Plag and Kunter's (2010) study informativeness was only tested via the calculation of the so-called constituent family size, whereas Bell and Plag (2012) also used other informativeness measures, such as the frequency or the semantic specificity of constituent N2, which emerged as significant predictors.⁷

⁷The constituent family size is the number of compounds that share the same left or right constituent with each other (cf. Bell 2008). For example, the compounds *summer house*, *summer night*, *summer dress* belong to the same left constituent family, which in this case consists of three members. It has been argued that compounds with a large left constituent family size tend to be right prominent,

To sum up, the review of the literature on prominence assignment in NN compounds has shown that contrary to Chomsky and Halle's (1968) CSR, one also finds a considerable number of right prominent NNs in English. Various different factors have been proposed in the literature in order to account for this variable prominence behaviour of NN compounds, yet not all of these factors also emerged as significant in experimental and corpus studies by Plag and his colleagues. According to these studies the constituent family bias, which functions as a predictor for analogy, seems to be the strongest determinant of compound stress. Furthermore, the studies revealed an independent effect of the factor semantics; compounds belonging to certain semantic subcategories, such as *N2 is located at N1* or *N2 is made of N1*, showed a greater tendency towards right prominence than compounds not belonging to these subcategories. Most recently, the study by Bell and Plag (2012) provides evidence for the assumption that informativeness affects prominence assignment in NN compounds, in that compounds with highly informative right constituents are more prone to right prominence than compounds with uninformative right-hand constituents.

Apart from providing important insights on the factors responsible for the variable prominence behaviour of compounds, the recent research literature on compound stress assignment also makes some contributions with respect to the phonological analysis of compound stress. In particular the studies by Kunter (2011) as well as Bell and Plag (2012) provide empirical support to an accent-based approach to compound stress as that proposed within the framework of autosegmental-metrical phonology (e.g. Ladd 1996).

According to this view, perceived prominence differences between the two constituents in a compound are marked by differences in the distribution of pitch accents (e.g. Ladd 1996:52; Gussenhoven 2004:18). Pitch accents are local features of the pitch contour, which indicate that the syllable with which they are associated is prominent in an utterance (cf. Ladd 1996:48).⁸ Potential landing marks for these pitch accents are only stressed syllables in English, whereas unstressed syllables do not provide such accentual targets. Thus, within this framework a difference is made between unstressed syllables on the one hand, and stressed syllables with and without a pitch

whereas compounds with a large right constituent family size favour left prominence (cf. Bell 2008). The idea behind this assumption is that the more likely a certain item occurs in the left or right position in a compound, the more it is expected in this position, and thus the less informative it is (cf. Plag and Kunter 2010).

⁸Pitch accents come in different shapes, i.e. a pitch accent may be realized as a high tonal target (H*) or a low tonal target (L*) or it may be a combination of the two types (e.g. H*+L; L+H*)(e.g. Pierrehumbert 1980).

accent on the other hand (cf. e.g. Gussenhoven 2004:20). Crucially, stressed syllables associated with a pitch accent are assumed to be perceived as more prominent by listeners than stressed syllables without a pitch accent (e.g. Ladd 1996).

Proponents of this approach have argued that prominence differences between the two constituents in an NN construction are marked by the presence or absence of a pitch accent on the stressed syllable of the right constituent of the NN structure (Gussenhoven 2004:18); left prominent NNs have a single pitch accent on the left constituent, with the right constituent being deaccented. In contrast, in right prominent NNs both constituents are assumed to be accented, with the second accent being perceived as more prominent than the first one. This claim has also found empirical support in a recent study by Kunter (2011) dealing with the phonetics and phonology of compound stress. Kunter found that compounds perceived as left prominent indeed usually have a single pitch accent on the left constituent, with no pitch accent on the right constituent, whereas in compounds perceived as right prominent both constituents are accented. The difference between left and right prominent compounds in Kunter's (2011) study was phonetically marked by differences in pitch and intensity on the right one of the two compound constituents; left prominent NN compounds had a high pitch and intensity value on the left constituent with clearly lower values on the right constituent. In contrast, right prominent NNs had high pitch and intensity values on both constituents.

Additional support for an accent-based approach to compound stress is also provided in Bell and Plag's (2012) study. As noted earlier in this section, they found that the degree of informativeness of the right constituent of a given compound affects the compound's stress pattern. According to Bell and Plag (2012), this finding provides evidence for the assumption that compound stress assignment is a matter of accentuation since informativeness effects are highly expected under such an accent-based approach, yet they are less expected under an approach that argues in terms of different degrees of lexical stresses (cf. Bell and Plag 2012:43).

Thus, there is strong empirical support for the assumption that in terms of phonology dealing with compound stress means that one deals with patterns of accentuation. In view of this situation, I will avoid the term 'stress' in the remaining of this thesis and instead prefer the term 'prominence'. The term prominence seems more suitable to capture the fact that pitch accents lend additional prominence to an already stressed syllable. Consequently, when referring to prosodic differences between the constituents of left- and right-branching compounds I will use the terms 'left prominent' and 'right prominent' rather than 'left stress' and 'right stress'.

2.2 Prominence assignment in triconstituent

Noun+Noun+Noun compounds

According to the generative approach by Chomsky and Halle (1968), prominence assignment in triconstituent Noun+Noun+Noun compounds is governed by the same rule that assigns prominence in Noun+Noun compounds, namely the ‘Compound Stress Rule’ (CSR). Due to its recursive nature and its cyclic application, the Compound Stress Rule assigns highest prominence to the leftmost constituent in a left-branching compound (e.g. [*dáy care*] *space*). For triconstituent compounds with a right-hand complex constituent (e.g. *area* [*stréet maps*]), however, it is generally assumed that highest prominence falls on the second constituent of the whole compound. Since the classical Compound Stress Rule would not predict this prominence pattern, Chomsky and Halle added a structural constraint to the Compound Stress Rule to derive the expected prominence pattern for right-branching compounds (Chomsky and Halle 1968:93, example 70).

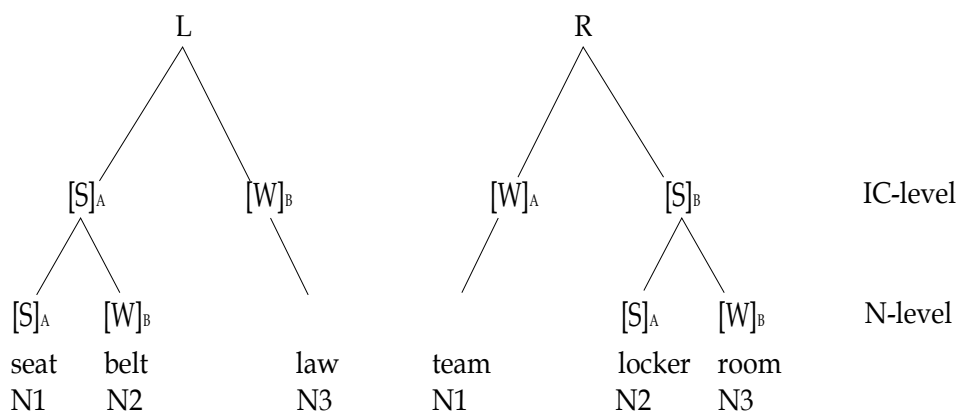
Liberman and Prince (1977) adopted Chomsky and Halle’s generalizations and incorporated them into their own theory of stress within the framework of metrical phonology. Their version of the modified Compound Stress Rule, the ‘Lexical Category Prominence Rule’, labels metrical trees on the basis of strong-weak relations between two sister constituents. It is assumed that one constituent is always strong (S), i.e. more prominent, with respect to its immediate sister constituent, which is automatically assigned a weak (W) status. The LCPR is thought to apply simultaneously at every level of the syntactic tree, which ensures that relative prominence is preserved under embedding (cf. Liberman and Prince 1977:256). The LCPR prediction as formulated by Liberman and Prince (1977) is given in (3) and illustrated again with the two examples shown in Figure 2.1.

- (3) “In a configuration [cA Bc]: if C is a lexical category, B is strong iff it branches” (Liberman and Prince 1977:257).

Figure 2.1 shows the metrical trees for the left-branching compound *seat belt law* and the right-branching compound *team locker room*.⁹ A look at Figure 2.1 shows that at

⁹The labels ‘N-level’ and ‘IC-level’ given in Figure 2.1 are not part of Liberman and Prince’s notation but were added by the author of the present thesis. The N-level marks the level at which the NN compound is created; the IC-level marks the level at which the triconstituent compound is created, namely by adding a third noun either to the left or the right of a previously constructed NN compound. At the IC-level of a compound, we deal with the prominence relation between the complex constituent as a whole and the single constituent.

Figure 2.1: Metrical tree of a left- and right-branching compound



the level of the complex constituent *seat belt* (N-level), the right-hand node B is not branching. Because of that, the LCPR assigns a strong label to the left constituent of the embedded NN compound and a weak label to the right constituent. The situation is similar at the immediate constituent level (IC-level) of the left-branching compound. Node B, which dominates the item *law*, is not branching which is why it is assigned a weak status by the rule. As a consequence, node A, which dominates the complex constituent on the left, is automatically labelled strong by the LCPR. The tree shows that the constituent *seat* is entirely dominated by strong nodes. This makes *seat* the so-called ‘designated terminal element’ in Liberman and Prince’s terminology, which is the most prominent constituent of the whole compound (cf. Liberman and Prince 1977:259).

In the right-branching compound *team locker room*, the prominence relation between the two constituents of the embedded NN compound is similar to that of the embedded NN compound in the left-branching compound *seat belt law*. The left constituent of the embedded NN compound *team locker* is assigned the strong label due to the absence of a right-hand branching node B at that level. Yet, at the IC-level of the compound the rule marks node B as strong because it is branching. Since *locker* is entirely dominated by strong nodes it becomes the most prominent constituent of the entire compound.

At first sight the LCPR, i.e. the factor branching direction, may seem relatively straightforward.¹⁰ However, there are some major problems associated with the LCPR and its predictions, which are discussed in more detail in the following paragraphs.

¹⁰The LCPR found its way into pertinent phonology and morphology textbooks such as Giegerich (1992), Plag (2003) and Spencer (1996).

First, a crucial shortcoming of the LCPR is that it is based on the assumption that compounds are categorically left prominent. This underlying assumption is captured in the “iff” expression of the rule. The ‘iff’ expression is to be paraphrased as ‘if and only if’ and guarantees that compounds, may they be NNs or left-branching NNNs, are always left prominent (cf. Liberman and Prince 1977:260). However, as mentioned in section 2.1, various studies (e.g. Plag et al. 2007; Bell 2008; Lappe and Plag 2007; Plag et al. 2008; Kunter 2011) dealing with prominence in biconstituent compounds have shown that apart from left prominent NNs one also finds a considerable number of right prominent NN compounds in English. The existence of these right prominent NN compounds on the one hand, and the assumption that prominence is preserved under embedding on the other hand, yet, may cause some trouble for the LCPR and its prediction that highest prominence is generally assigned to constituent N1 in left-branching compounds and constituent N2 in right-branching compounds. Instead, it is reasonable to believe that one may also find right prominent NN compounds embedded in triconstituent NNN compounds which may cause highest prominence on constituent N2 in left-branching compounds and on constituent N3 in right-branching compounds, *contra* the LCPR.

In fact, some first hints towards this assumption are provided by Olsen (2000), Giegerich (2009) and surprisingly even by Liberman and Prince (1977) themselves. Within a general discussion about the status of right prominent NN constructions, Olsen (2000:65) argues that both left and right prominent NN compounds occur embedded in triconstituent compounds (e.g. [*silicon chip*] *manufacturer*, [*oval office*] *visit*, [*time Warner*] *company*). Unfortunately, Olsen provides no examples for right-branching NNN compounds with embedded right prominent NN compounds. Yet, such examples are found in a recent theoretically based approach by Giegerich (2009), in which he argues that the LCPR and its predictions are wrong. Giegerich (2009:10) provides examples of both left- and right-branching triconstituent compounds with embedded right prominent NN compounds and argues that in these compounds highest prominence is assigned to the right constituent of the complex constituent instead of the left one (e.g. [*toy car*] *collection*, [*school office*] *manager*, *aluminium* [*garden shed*], *university* [*spring term*]).

Finally, Liberman and Prince (1977) themselves mention that right prominent NN compounds may occur embedded in triconstituent compounds and cause a prominence pattern different from that normally predicted by the LCPR. In particular, they argue that it is theoretically possible that a right prominent NN compound such as *Madison Avenue* may be embedded in a right-branching NNN compound, which causes

constituent N3 to be the most prominent constituent of the whole compound. However, Liberman and Prince (1977) also point out that such right-branching NNN compounds with embedded right prominent NNs are usually impossible to construct in English since right prominent NNs are phrases and compounds consist only of lexical constituents (cf. Liberman and Prince 1977:261).¹¹ Yet, irrespective of whether such compounds are rare or not, the crucial point is that the example provided by Liberman and Prince shows that the existence of right prominent NNs may have serious consequences for prominence assignment in triconstituent compounds.

Second, since right prominent NN compounds provide evidence that node B can be strong although it is not branching, this fact might also have serious implications for prominence assignment at the IC-level of triconstituent NNN compounds. In particular, one would not only expect to find left-branching compounds with highest prominence on constituent N1 but also a number of left-branching compounds with highest prominence on constituent N3. Such left-branching NNNs with highest prominence on constituent N3 are indeed documented in the literature (e.g. Kingdon 1958; Fudge 1984; Hayes 1995; Sproat 1994; Giegerich 2009) and again even by Liberman and Prince (1977) themselves (e.g. *whale-oil lamp*). However, most of these authors including Liberman and Prince do not refer to such left-branching NNNs as being exceptions to the LCPR. Instead, they refer to them as being phrases, whose prominence pattern is governed by the rule given in (4).¹²

- (4) “In a configuration [cA Bc]: If C is a phrasal category, B is strong” (Liberman and Prince 1977: 257).

Yet, the problem with this approach to explain away apparent exceptions to the LCPR is that apart from prominence itself no other independent criteria are provided by these authors that would distinguish NNN compounds from NNN phrases. This situation at the IC-level of left-branching NNN compounds is analogous to the situation described for right prominent NN compounds in section 2.1. By arguing that left-branching NNN compounds with highest prominence on constituent N3 are phrases, the LCPR and its predictions are rescued. In contrast, if one does not draw this rather arbitrary distinction between compounds and phrases, the LCPR fails to predict the

¹¹The example is only mentioned by Liberman and Prince (1977) in order to illustrate that the LCPR operates independently at every level of the tree, a property that distinguishes their formulation of the Compound Stress Rule from that of Chomsky and Halle (1968).

¹²The LCPR predicts, as does the Compound Stress Rule, prominence only in lexical categories. For prominence assignment in phrasal categories, Liberman and Prince (1977) formulate a rule equivalent to Chomsky and Halle’s (1968) ‘Nuclear Stress Rule’ (NSR) which assigns highest prominence to the rightmost constituent in a phrase.

correct prominence pattern for these NNN compounds. However, if it is not the phrasal status of these left-branching NNNs that triggers highest prominence on constituent N3, what else may trigger this prominence behaviour?

Under the assumption that triconstituent compounds are also binary in nature (e.g. Plag 2003:134), one possible answer to this question might be that the same factors that trigger right prominence in NN compounds (N-level) may also trigger right prominence at the IC-level of left-branching NNN compounds; an assumption that has also been recently put forward by Giegerich (2009). As discussed in section 2.1 the factors that were claimed to be responsible for prominence variation in NN compounds were semantics, analogy, structure and informativeness of which, however, only semantics, analogy and most recently informativeness also emerged as significant. Hence, it seems reasonable to believe that semantics, analogy and informativeness may also govern prominence assignment in left-branching NNN compounds. Some first hints for the potential influence of semantics on prominence assignment to triconstituent compounds can even be found in the literature. In a study dealing with prominence assignment in NN and NNN constructions, Sproat (1994) argues that the same semantic phrase structure rules that trigger right prominence in NN constructions also trigger right prominence in NNN constructions.¹³

In addition to these theoretical shortcomings of the LCPR, a third problem is that the LCPR is only poorly empirically supported. The literature on prominence assignment in triconstituent compounds is scarce and the LCPR predictions are primarily illustrated by the same self-selected examples repeated throughout the literature. The only study I am aware of which might be considered as providing empirical proof for the LCPR and its prediction is the one by Sproat (1994). However, Sproat (1994) takes the LCPR for granted and applies it to his (written) data, instead of actually testing it. In addition, the results are also based on an arbitrary and problematic assignment of phrasal vs. compound status to NNN structures. For instance, Sproat (1994:82) assigns compound status to the NNN sequence *sump pump factory* whose prominence pattern is then governed by the LCPR, whereas *living room table* is assigned phrasal status based on the assumption that a semantic phrase structure rule such as “room + furniture” marks phrases. Due to being a phrase, *living room table* exhibits highest prominence on the constituent *table*. However, ‘room + furniture’ structures are again only regarded as being phrases due to their right prominent patterns, which renders the whole approach circular.

¹³For more details on this study see the following paragraph.

Besides the study by Sproat (1994), only a few other studies are available that deal with prominence assignment in NNN compounds, and their results indicate that the LCPR is problematic in its empirical predictions. For instance, Kvam (1990) investigated 40 Noun+Noun+Noun constructions in a production experiment. Kvam found that the majority of the investigated compounds, namely 30 out of 40, was either exclusively or by the majority of the experimental subjects stressed on constituent N2. Yet, Kvam points out that only 10 of these compounds were also clearly right-branching, i.e. the group of compounds that should indeed have this prominence pattern. Hence, only in 10 cases prominence assignment could be directly related to the branching direction of the compound. Based on his findings, Kvam concludes that prominence does not necessarily serve to indicate constituency structure but that it is primarily a means of emphasis rather than basic meaning (Kvam 1990:158). Unfortunately, Kvam does not provide any information on the criteria on which he based the selection of his test items nor does he explicitly state how he detected the prominence pattern of a given compound. Finally, Kvam does not mention which of the compounds under investigation were prominent on which constituent, which would be necessary for a more thorough investigation of the problem.

Apart from Kvam's study, additional evidence towards more variation in the prominence assignment of NNN constructions is provided by Berg (2009). Taking an explorative approach by looking at a total of 642 Noun-Noun-Noun combinations taken from the BNC, Berg (2009:87) finds that 57.2% of the combinations have highest prominence on constituent N2, and 26.5% on N1. Thus, Berg's findings go in the same direction as Kvam's results, revealing a general tendency for triconstituent compounds to exhibit in their majority highest prominence on the second constituent, be they left- or right-branching. However, this tendency is statistically more significant for right-branching than for left-branching compounds, which is a finding that is more in line with the LCPR. In addition to that, Berg (2009:88) also provides information about a number of right-branching compounds with highest prominence on constituent N1 and N3, as well as left-branching compounds with highest prominence on N3. With reference to the LCPR prediction all of these compounds would be considered violations, either at the N-level or at the IC-level, although Berg does not explicitly refer to the LCPR and its predictions.

However, the assignment of the prominence pattern to left- and right-branching compounds in Berg's study is based on the author's own intuition about prominence as well as on the intuition of a few other judges (personal communication, November 2008). This method must be considered problematic, if applied impressionistically

only. As studies regarding prominence assignment in NN compounds have shown, listeners tend to vary in their judgements (e.g. Bauer 1983a:103; Kunter 2011). Therefore, assigning prominence to triconstituent compounds solely based on one's own intuition should be avoided and replaced by a more objective method, for example by a controlled rating procedure or by using measurements of the acoustic correlates of prominence (e.g. Plag et al. 2008).

The review of the literature on prominence assignment in triconstituent nominal compounds has shown that the LCPR is highly questionable from a theoretical as well as from an empirical point of view. On the one hand the existence of right prominent NN compounds is ignored in the formulation of the rule, which may have serious consequences for prominence assignment in triconstituent compounds both at the N-level and the IC-level. On the other hand, empirical investigations of the rule are scarce and there are serious implications that the rule is far from being categorical in its predictions. Given this situation, the current thesis addresses the following research question and new hypotheses regarding prominence assignment in triconstituent NNN compounds.

1. To what extent does the LCPR predict the correct prominence pattern of left- and right-branching NNN compounds in English?

2. **Embedded Prominence Hypothesis (EPH):**

If highest prominence falls on the complex constituent of a triconstituent compound:

a) left- and right-branching compounds with an embedded left prominent NN compound have highest prominence on the left member of the complex constituent, i.e. N1 and N2, respectively.

b) left- and right-branching compounds with an embedded right prominent NN compound have highest prominence on the right member of the complex constituent, i.e. N2 and N3, respectively.

3. **IC-Prominence Hypothesis (IPH):** The same factors thought to trigger right prominence in biconstituent NN compounds, i.e. semantic and analogical factors, also trigger right prominence at the IC-level of left-branching NNN compounds.

The 'Embedded Prominence Hypothesis' given in 2 predicts prominence for left- and right-branching compounds with both embedded left and right prominent NN compounds. Yet, as implied by the 'if' condition of the hypothesis the predictions of the

EPH focus only on the embedded NN and its potential effect on the prominence behaviour of triconstituent compounds. Hence, the EPH does not want to say anything about the factors operating at the IC-level of triconstituent compounds, but assumes that highest prominence falls on one of the two members of the complex constituent. Part a) of the EPH basically makes the same predictions as the LCPR, in that it predicts the left member of the embedded NN to be the most prominent constituent of the whole compound. In contrast, part b) makes predictions for left- and right-branching compounds with embedded right prominent NNs. These compounds are predicted to have highest prominence on the right member of the complex constituent, a prominence pattern not predicted by the LCPR.

The 'IC-Prominence Hypothesis' given in 3 predicts, as the name implies, prominence relations at the IC-level of triconstituent compounds. In particular, the IPH predicts that semantics and analogy trigger right prominence at the IC-level of left-branching compounds (i.e. highest prominence on N3), overwriting the factor branching-direction. The IC-Prominence Hypothesis is restricted to left-branching compounds in this thesis as I decided to focus first of all on the question whether the same factors triggering right prominence in biconstituent compounds, i.e. at the N-level, also trigger that prominence pattern at the IC-level of triconstituent compounds.

In addition, the IC-Prominence hypothesis concentrates on semantics and analogy in this thesis. The factor 'informativeness', which has also been discussed in this section, is not considered here as it has only very recently been found to influence prominence assignment in NN compounds (e.g. Bell and Plag 2012).

The research question given in 1 is addressed in a corpus study. In contrast the Embedded Prominence Hypothesis and the IC-Prominence Hypothesis are tested by means of two production experiments conducted with native speakers of North American English at the University of Toronto. The choice for an experimental design in order to test the potential influence of right prominent NN compounds on prominence assignment in triconstituent compounds allowed me to explicitly control for relevant factors such as the prominence pattern of the embedded NN compounds or the semantic relations at the IC-level of the triconstituent compounds, i.e. the semantic relation between the single and the complex constituent in a compound.

Having introduced the research question and hypotheses addressed in this thesis, the next chapter deals with the general methodology used in order to determine the prominence patterns of the triconstituent NNN compounds under investigation.

3 General Methodology

3.1 How to determine prominence patterns of English compounds?

Investigating prominence assignment in English triconstituent Noun+Noun+Noun compounds raises the crucial question of how to determine whether a left-branching compound has highest prominence on constituent N1 and a right-branching compound on constituent N2. As already mentioned in section 1, thus far the generalizations about prominence assignment in triconstituent compounds rely primarily on the researchers' own intuition about prominence (e.g. Liberman and Prince 1977; Liberman and Sproat 1992; Sproat 1994; Berg 2009; Giegerich 2009). Yet, this way of assigning prominence to a given compound is highly problematic. As noted in various studies dealing with prominence assignment to Noun+Noun compounds (e.g. Bauer 1983b; Plag 2006; Kunter and Plag 2007; Kunter 2011), speakers, experts they may be, often seem to have difficulties in classifying compounds as either left or right prominent and thus do not only vary within their own prominence judgements but also among each other. Because of that, Bauer (1983b:51) notes that prominence patterns of compounds derived by introspection should be regarded as untrustworthy. Given this problem associated with individual speaker judgements, one would like to have a more systematic approach to determine the prominence patterns of triconstituent compounds.

A more systematic and reliable approach may be that of a controlled rating procedure with a very large number of listeners. In particular, one may conduct an experiment in which a large group of listeners is asked to say whether compounds exhibit highest prominence on constituent N1, constituent N2 or constituent N3 and in which the reliability of the raters is carefully controlled for. Such an experiment was for instance conducted by Kunter (2011) as part of a larger study investigating the acoustic correlates of compound prominence and testing hypotheses about compound prominence assignment in biconstituent Noun+Noun compounds. In his ex-

periment, Kunter extracted 105 compounds from the Boston University Radio Speech Corpus (Ostendorf et al. 1996) and presented these compounds on a computer screen to 32 participants. At first, the participants were asked to listen to the compounds and then to use a slider to indicate on a scale, ranging from 0 to 999, as to whether they perceived a given compound as left prominent, right prominent or with equal prominence on both constituents. The more the participants perceived a compound as right prominent, the more they moved the slider to the right, the more the compound was perceived as left prominent, the more the slider was moved to the left. Finally, when the slider was moved to the middle of the scale it indicated level prominence, i.e. the listener had no clear preference for either left or right prominence, but perceived the two constituents as equally prominent.

Crucially, Kunter was able to account for the reliability of his raters. By applying different statistical tools, he determined a sub-group of 17 raters who turned out to be fairly consistent in their ratings. In particular, the ratings of these listeners corresponded to the general trend and showed only small within-group variation. Given these consistent ratings, Kunter classified these raters as proficient raters. Furthermore, he determined 12 raters whose ratings turned out to be less reliable in the sense that these raters either showed a tendency to use only the middle of the rating scale for both left and right prominent compounds, or they showed a larger random variation around the general trend than the 17 proficient raters (cf. Kunter 2011:53). Finally, Kunter determined two raters that differed considerably in their ratings from all the other raters and one rater that seemed to have failed to follow the experimental instructions. These participants were classified as non-proficient raters. Being able to differentiate between proficient listeners on the one hand, and less proficient and non-proficient listeners, respectively, on the other hand, Kunter was able to select only those ratings for further analysis that came from the reliable raters, which increased the validity of his results.

Kunter's study illustrates how statistical measures may be applied to account for the reliability of listener ratings and thus to increase the validity of one's perception data. However, conducting such a perception experiment in order to determine prominence assignment in triconstituent NNN compounds might still not be an appropriate choice. Given that speakers already seem to have difficulties in classifying Noun+Noun compounds as either left or right prominent, they may have even more difficulties when confronted with the task of rating prominence patterns of triconstituent compounds. Thus, the number of listeners that may fail to manage such a task or show considerable variation in their ratings may be expected to be relatively

high. Consequently, when conducting such an experiment with triconstituent compounds one may need a very large number of listeners in order to avoid running out of enough valid data.

Hence, an alternative and more objective approach towards determining prominence assignment in English compounds would be that of measuring the acoustic correlates of compound prominence as done in various experimental and corpus studies by Farnetani et al. (1988), Plag (2006), Kunter and Plag (2007), Plag et al. (2008) and Kunter (2011). According to those studies prominence differences in compounds are phonetically marked by differences in the fundamental frequency (F0), intensity, duration and spectral balance. The fundamental frequency is the rate of vocal cord vibration and is generally argued to be the acoustic correlate of what listeners perceive as pitch. Intensity is the acoustic correlate of what listeners perceive as loudness, whereas 'duration' refers to the phonetic correlate of length (e.g. Ladefoged 2003). Finally, the spectral balance captures intensity differences across the entire frequency spectrum. Yet, although all of these acoustic cues were found to be relevant in order to distinguish between left and right prominent NNs, the above mentioned studies revealed that pitch seems to be the strongest predictor to prominence, with the effects for the other acoustic cues being also significant but to a lesser degree.

For instance, Farnetani et al. (1988) investigated minimal pairs such as *páper bag* and *paper bág* in order to detect the acoustic cues responsible for the different prominence patterns of phrases and compounds. They measured pitch and intensity peaks as well as duration in the left and right constituent of their data and statistically compared the values obtained for the compounds to those of the corresponding phrases. They found that the difference in pitch and intensity between the left and the right element was larger for the NN compounds than for the corresponding NN phrases of the minimal pairs, yet with the effect for pitch being stronger than for intensity. With reference to duration, Farnetani et al. (1988) found that this cue only played a role in the distinction between members of minimal pairs, i.e. members of phrases were longer than those of compounds. Hence, for a distinction between left and right prominence this cue seemed insignificant.

A similar result to that of Farnetani et al. (1988) is reported in the study by Ingram et al. (2003), who were also interested in the acoustic cues responsible for the different prominence patterns of compounds and phrases. In addition to compounds and syntactic phrases, Ingram et al. (2003) also analysed syntactic phrases with a contrastive focus (e.g. *it was a **black** berry, not a **red** one*). In their study Ingram et al. (2003) noted that the pitch change between the left and right constituent in the compounds and

the contrastive phrases was statistically significantly larger than the pitch change in the corresponding syntactic phrases. Based on their result, Ingram et al. (2003) concluded that the most important cue to prominence was pitch, with intensity playing only a supportive role. In contrast, duration was again only found to be crucial with reference to the distinction between members of minimal pairs.

The importance of the F0 as an acoustic cue to prominence was also determined in Kunter's (2011) study dealing exclusively with prominence assignment in NN compounds. In one of a series of different analyses, Kunter addressed the question in how far the perceived prominence of compounds is interrelated with the phonetic properties of the F0, intensity, duration, pitch slope and spectral balance. Kunter (2011) measured those 5 acoustic cues in the primarily stressed syllable in the left and right constituent of 105 NN compounds, which he had randomly selected from the Boston University Radio Speech Corpus. These compounds also provided the data base for a previously conducted perception rating by Kunter. Because of that, for each compound Kunter had an average value for each of the five acoustic correlates of prominence on the one hand, and perception scores on the other hand. He found that compounds perceived as left prominent are characterized by a higher pitch and higher intensity on the left member and a clearly lower pitch and intensity on the right member. In contrast, compounds perceived as right prominent showed indistinguishable values for F0 and intensity on both constituents (cf. Kunter 2011:93). Furthermore, Kunter (2011) found that in right prominent compounds the primarily stressed syllable of the left constituent is shorter than the primarily stressed syllable of the right constituent while in left prominent NNs the two syllables are of equal length. Crucially, the effects for pitch and intensity were the strongest according to Kunter's analysis.

The dominant role of the F0 was also determined in a slightly different analysis by Kunter (2011).¹ In this analysis Kunter tried to predict the prominence ratings obtained in the above mentioned perception experiment on the basis of acoustic measurements of pitch, intensity, duration and spectral balance. He devised a regression model with the different acoustic cues as model parameters and with the mean perception ratings for each compound as dependent variable. Kunter's regression model turned out to be quite successful in predicting the perception scores of his rating experiment. Crucially, a closer investigation of the partial effects of the regression model revealed that the effect for pitch was the strongest in the model (cf. Kunter 2011:114f).

Finally, an experimental study by Plag (2006), in which he analysed about 500 com-

¹See also Kunter and Plag (2007) for an earlier version of this analysis.

pounds, showed that measuring pitch and calculating pitch differences between the left and right member of a compound is quite suitable to test competing hypotheses regarding prominence variation in NN compounds. Plag measured pitch in the most prominent syllable in each compound constituent and compared the pitch behaviour of compounds which were assumed to be left prominent (e.g. argument-head compounds) with those claimed to favour right prominence (e.g. modifier-head compounds). Plag indeed found statistically significant, and expected, differences in the pitch behaviour between these groups, i.e. the pitch change between the left and right constituent in argument-head compounds was generally larger than the pitch change in modifier-head compounds. On the basis of his findings, Plag was able to make clear statements with reference to the accurateness of the hypotheses tested (cf. Kösling and Plag 2009:211f).

Given the potential problems that speakers may encounter when having to rate the prominence patterns of triconstituent compounds on the one hand, and given the robust findings for the F0 obtained in studies using measurements of the acoustic correlates of prominence on the other hand, I decided to measure the F0 in order to determine the prominence patterns of triconstituent NNN compounds. The measurement of the F0 was done automatically by using the same pitch algorithm as the one used by Kunter (2011) (For more details on how the pitch algorithm works the reader is referred to section 4.1 of this thesis and Kunter (2011), respectively.). Measuring the F0, however, also involves a number of methodological problems which are discussed in more detail in the following section.²

3.2 Pitch measurements

The first methodological problem associated with measuring the F0 as an acoustic cue to prominence is that of finding an appropriate measurement point for it. In the respective literature one finds three different measurement points; the F0 in the middle of the vowel, the mean F0 over the sonorant part of the rime of the highest prominent syllable and the F0 peak in the highest prominent syllable. I will discuss each of these measures in turn, beginning with the middle of the vowel as chosen in studies by Ingram et al. (2003) and Plag (2006).

In their studies, Ingram et al. (2003) and Plag (2006) decided to measure the F0 value in the middle of the vowel of the highest prominent syllable in each constituent

²In the remaining of this thesis I will use the terms pitch and fundamental frequency (F0) more or less interchangeably.

and then calculated pitch differences between the left and right constituent of each compound. In choosing this point of measurement they tried to account for so-called coarticulation effects caused by surrounding sounds (cf. Kösling and Plag 2009). For instance, Plag (2006) notes that vowels that occur in a voiceless environment tend to exhibit a higher pitch than vowels surrounded by voiced sounds and that this effect is particularly strong in the transition from consonant to vowels (cf. Plag 2006:150). An example of such a coarticulation effect is provided by Ladefoged (2003:87), who notes that

[...] there is a considerable difference between the beginning of *buy* (which is virtually level) and the beginning of *pie* (which descends rapidly). The high pitch at the beginning of *pie* is due to the high rate of airflow for the **ph**, which continues into the beginning of the vowel, producing a higher rate of vibration of the vocal folds. (Ladefoged 2003:87)

However, choosing the middle of the vowel as a measurement point for pitch is also associated with some problems. For instance, Ingram et al. (2003) and Plag (2006) had to deal with the technical problem of clearly separating the vocalic nucleus from preceding and following sounds, which is especially difficult in cases of approximants ([l],[r],[j],[w]). These sonorous sounds show similar acoustic properties to those of vowels, which makes a clear separation of these sounds often impossible (cf. e.g. Ladefoged 2003:98). In addition to that, Kunter (2011:75) points to another disadvantage associated with this measurement point. He notes that by measuring pitch in the middle of the vowel, one assumes that the middle is representative for the entire pitch contour of the segment. This assumption, however, is problematic given the existence of complex pitch accents such as L-H* (low-high). In these accents the high target is not necessarily found in the middle of the vowel, but may also occur later in the syllable or even in the beginning of the following unstressed syllable (cf. Ladd 1996:55). Hence, by measuring pitch in the middle of the vowel, one may fail to account for such rising accents, which may obscure the result of the pitch analysis.

A potentially more adequate way of measuring pitch was used by Kunter and Plag (2007), Plag et al. (2008) and Kunter (2011), who, instead of measuring the F0 value in the middle of the vowel, calculated the mean F0 over the sonorous part of the rime in the pertinent syllable (cf. Kösling and Plag 2009:212). Choosing the rime of the highest prominent syllable as measurement interval solved the above mentioned problem of having to separate the nucleus from following sounds. Furthermore, measuring the mean F0 over the sonorant part of the rime also takes care of the problem that accen-

tual tonal targets need not be the vowels themselves (cf. Ladd 1996:55). Finally, the F0 means also allow for a manual checking and recalculation of the pitch values returned by the pitch algorithm. As noted by Kunter (2011:75), this becomes particularly useful if the measurement interval is affected by non-modal phonation such as creaky voice. In creaky voice the rate of vocal cord vibration is much lower than in modal phonation (e.g. Ladefoged 2006), which may cause the algorithm to return either wrong or no pitch values for affected segments. Hence, a subsequent manual calculation of the mean pitch becomes necessary.

Finally, instead of using the F0 mean, one might also consider measuring the peaks of the pitch contour (e.g. Farnetani et al. 1988). However, Kunter (2011), who measured both the pitch peak and the mean F0, showed that the values of F0 peaks and F0 mean strongly correlated for his data, so that both measurements seem in principle suitable to account for prominence difference between left and right prominent compounds. However, the pitch peak as a measurement point has a major disadvantage in relation to the mean F0. Measuring pitch peaks in automatic pitch tracking runs the danger that the algorithm confounds peaks associated with so-called boundary tones with accentual tones (cf. Kunter 2011:74). Such boundary tones are intonational tones which appear only at the edges of prosodic constituents, but are not aligned with stressed syllables to indicate phonological prominence (cf. Gussenhoven 2004:22). Hence, a confusion of boundary tones with pitch accents should be avoided.

Thus, given these advantages the mean F0 has over the other possible measurement points, I decided to follow Kunter and Plag (2007), Plag et al. (2008) and Kunter (2011) in choosing the mean F0, calculated over the sonorant part of the rime of the highest prominent syllable of each compound constituent.

A second problem associated with acoustic measurements of pitch occurs with reference to the intrinsic pitch of vowels. It has been noted that high vowels (e.g. [i],[u]) have a higher intrinsic pitch than low vowels (e.g. [a]) (e.g. House and Fairbanks 1953; Lehiste 1970; Whalen and Levitt 1995). Given this fact, an analysis which uses pitch measurements to account for prominence differences between different types of compounds, should try to control for these vowel intrinsic pitch differences in order to prevent them from obscuring the outcome of the analysis (see also Gussenhoven 2004:9 on this point). In particular, if one does not control for those differences, it may be possible that significant pitch differences between different types of compounds are not the result of prominence differences, but are triggered by the different vowels in the pertinent syllables. Similarly, it is possible that one does not find expected pitch differences, i.e. prominence differences, between different types of compounds as they

are obscured by differences in the intrinsic pitch of vowels. But how can this factor be controlled for?

With reference to the experimental data of this thesis one could have tried to construct well-balanced stimuli in order to neutralize these differences. In particular, one could have construct triconstituent compounds in which the vowel of the highest prominent syllable in each constituent is held constant within a given compound as well as across different compounds. However, the construction of such stimuli for the two production experiments of the current thesis turned out to be impossible as other factors, crucial for testing the different hypotheses, had to be controlled for as well; for example the branching direction of the compounds, the prominence pattern of the complex constituents and the semantic relation between the complex and single constituent of the compounds. In addition to that, the construction of such stimuli would not have solved the same problem for the corpus data, for which the construction of specific stimuli had never been a viable option in the first place.

Yet, another method that seemed suitable for both the corpus and the experimental data was that of statistically controlling for vowel intrinsic pitch differences. For instance, in a study by Plag et al. (2011) dealing with the acoustic correlates of primary vs. secondary stresses in polysyllabic words, it is reported that the influence of intrinsic pitch differences may be at least partially accounted for by modelling the stimuli as random effects factor in a so-called mixed-effects model. Random effects factors of a statistical model are thought to account for the variance introduced by factors which are not of primary interest to the research question at hand, as in our case vowel intrinsic pitch differences.³ Plag et al. (2011) included the vowels of the relevant measurement intervals as covariate into a mixed-effects model and fitted one model with the stimuli modelled as random effect and one model without the stimuli included as random effect. They found that the F₀ effect for the vowels was significant in the model without the stimuli as random effect but non-significant in the model with the stimuli as random effect. Their result indicates that modelling the items of the experiment as random effects in a mixed-effects model helps to alleviate the influence of intrinsic pitch differences (cf. Plag et al. 2011:366).

The experimental and the corpus data of the present thesis were also modelled by means of linear mixed-effects models with the items fitted as random effects. Because of that, I am confident that the potential influence of vowel intrinsic pitch differences has been largely reduced in the different studies. In addition to that, I believe that the decision to measure the mean F₀ over the sonorant part of the rime additionally helps

³For more details on random effects and mixed-effects models see section 3.4.

to reduce potential effects of vowel intrinsic pitch differences. The reason for this is that in many constituents the highest prominent syllable also contains diphthongs, or vowels accompanied by sonorant material over which the mean is calculated, too. This may further diminish the effect (cf. Kösling and Plag 2009:212)

A third problem associated with pitch measurements that needs to be discussed here is the observation that pitch generally declines over the course of an utterance (e.g. Collier 1975; Pierrehumbert 1979). The reason for this pitch declination effect is seen in the lowering of the subglottal pressure during exhalation phases in the speech production process. This lowering of the subglottal pressure results in a lower rate of vocal cord vibrations and thus in a decrease of the fundamental frequency, i.e. pitch (cf. Gussenhoven 2004:97).

Studies by Plag (2006), Plag et al. (2008), Kunter and Plag (2007) and Kunter (2011) have shown that this pitch declination effect is also found between the constituents of compounds, in that pitch generally decreases from the left to the right constituent of a compound, may it be left or right prominent. Hence, given this effect, right prominent compounds do not necessarily have a higher pitch on their right constituent than on their left constituent. Instead, these studies revealed that right prominent compounds are typically marked by having either equally high pitches in both constituents or even a slightly lower pitch on the right-hand constituent. Crucially, with reference to the perception of prominence it has been shown that listeners are able to make up for this pitch declination effect, still perceiving compounds as right prominent even if the right member has a slightly lower pitch than the left member (e.g. Kunter 2011).

However, if one only measures pitch in order to determine the prominence patterns of compounds, as it is done in the present thesis, this natural decline of the F0 raises the question of how to determine whether a given compound is left or right prominent? At which point can we say that a lower pitch on the right one of two compound constituents indicates left prominence? Similarly, how do we know that the prominence relation between two constituents is that of right prominence?

With reference to that, the studies by Plag (2006), Kunter and Plag (2007), Plag et al. (2008) and Kunter (2011) have shown that irrespective of this general pitch declination effect, one still finds statistically significant differences in the degree to which pitch drops between the two constituents of left- and right-prominent compounds. All of these studies showed that in right prominent compounds the pitch drop between the left and right constituent was statistically significantly smaller than the pitch drop between the two members in left prominent compounds. Given this result for biconstituent NN compounds, we would also expect to find such statistically significant

differences between the pitch values of triconstituent left- and right-branching compounds. Yet, in what way the pitch values of the left- and right-branching compounds should differ from each other in order for the LCPR and the two hypotheses (see section 2.2 again) to be true will be described in more detail in the pertinent chapters presenting the corpus study and the two experiments.

3.3 Discourse factors

Apart from problems directly associated with measuring pitch as an acoustic correlate of prominence, another problem arises with reference to analysing natural speech data. Using such data involves the problem of potential influences of discourse factors, such as contrastive stress, focus and the given/new distinction, on the prominence behaviour of compounds (see also Plag et al. 2008 for a discussion of this point). For instance, it has been argued that items representing discourse old information often tend to be deaccented by speakers (e.g. Hirschberg 2002). Furthermore, items which would normally not be marked as prominent in an utterance may be emphasized for contrastive purposes (e.g. *I want **this** apple not that one*). Such contrastive stress environments may also change the prominence pattern of compounds. Plag et al. (2008:772) for instance illustrate this effect with the example of *Park Street*. They point out that the compound *Park Street* is usually assumed to be left prominent, yet in a context like *I said Park Street, not Park Avenue* special emphasis may be placed on *Street* in order to contrast it with *Avenue* (cf. Plag et al. 2008:772). It is reasonable to believe that a similar effect may also be observed for triconstituent compounds occurring in contrastive stress environments; left-branching compounds may have highest prominence on the second and the third constituent, respectively, instead of the first one, as illustrated with the examples in (1). Furthermore, in right-branching compounds highest prominence may be assigned to constituent N1 or constituent N3 if the compound is placed in such a context (see (2)).

- (1) *He read about a coffee **table** designer, not a coffee **mug** designer.
He looked for Fish market **Street**, not Fish market **Avenue**.*
- (2) *She visited a **China** information center, not a **tourist** information center.
She was eager to point out that she had only listened to the radio morning **news** this morning, not the radio **celebrity** talk.*

Apart from contrastive stress, however, other discourse factors do not seem decisive for prominence assignment in compounds. For instance, Plag (2006) tested the effects of focus and the given/new distinction in that he embedded the compounds in different sentence positions and constructed different clause types. Although he found that the pitch values of the left and right constituents in each compound generally decreased from initial to final clausal position, no clear effects of these factors on prominence assignment could be detected. Furthermore, Plag et al. (2008) tested three hypotheses regarding prominence variation in NN compounds against a large number of corpus data from the Boston University Radio Speech Corpus. They found, among other things, that argument-head compounds are not generally more left prominent than modifier-head compounds. Instead, this effect is restricted to argument-head compounds with a deverbal head ending in *-er*. In order to rule out potential influences of discourse factors on their results, Plag et al. (2008) compared their findings to those of a study by Plag et al. (2007), in which the same three hypotheses were tested against (mainly) dictionary data taken from the CELEX lexical data base (Baayen et al. 1995). The CELEX study revealed a similar effect for argument-head compounds ending in *-er* and also revealed quite similar results with respect to the other two hypotheses. Hence no relevant differences in prominence assignment between citation forms and speech corpus data were found (cf. Kösling and Plag 2009:213).

In view of this situation, I checked the corpus data for occurrences of contrastive stress environments, but found no such environments among the data. Since the other factors do not seem to be decisive according to previous studies and given that the compounds selected from the corpus were randomly sampled, potential effects caused by discourse factors were neglected in the corpus study. Furthermore, when constructing the carrier sentences for the experimental data, I avoided to create contrastive stress environments. In addition, in order to entirely rule out potential effects of information structure on the prominence pattern of the experimental data, I decided to present all compounds as new information to the participants. Finally, I kept the sentential position of the compounds constant in all sentences (more details on the construction of the carrier sentences are given in the pertinent sections dealing with the experimental data).

3.4 Mixed-effects models

As noted in section 3.2, the data in the present thesis was statistically analysed by means of linear mixed-effects models. A linear mixed-effects model is a statistical

model that allows the inclusion of both fixed effects factors and random effects factors (e.g. Pinheiro and Bates 2000; Crawley 2005b; Baayen 2008). The possibility to include parameters for both fixed effects and random effects distinguishes mixed-effects models from ordinary regression models and Analysis of Variance (ANOVA), respectively, which only allow the inclusion of parameters for fixed effects. But, what is gained by adding additional parameters for so-called random effects to a statistical model? Why are mixed-effects models appropriate to analyse the data in the present thesis? In order to answer this question, it is necessary to briefly explain what is understood by fixed effects factors and random effects factors.

Fixed effects factors are variables that have a fixed number of repeatable factor levels (cf. Baayen 2008:241). It is often the experimental treatment factors, i.e. the variables that are of primary interest to the research question at hand, which are modelled as fixed effects factor in a statistical model (cf. Crawley 2005b:178). The levels of these factors represent different conditions, which are selected by the researcher on purpose and for which he seeks to find an effect on a given response variable. These treatments are repeatable in the sense of that they can be applied over and over again to any new data. For instance, a central aim of the present thesis is to find out as to whether the speakers' assignment of pitch to a given compound depends on the branching direction of that compound. Thus, in a statistical model, the response variable would be pitch and the branching direction of the compounds would be a fixed factor, with the two factor levels left- and right-branching. According to the LCPR, left-branching compounds should exhibit a different pitch behaviour than right-branching compounds. Hence, variation in pitch is expected to be associated with variation in the branching direction of a compound. Crucially, the two factor levels of the factor branching direction are repeatable, as one can always add new compounds to a given data set and code these compounds as either left- or right-branching.

In contrast to fixed effects factors, random effects factors do not have a fixed number of repeatable and thus controllable factor levels, but their number of levels is said to be indefinite (cf. Baayen 2008:241). The reason for this is that random factors are factors with levels randomly sampled from a larger population and replicating a given study would involve choosing another sample from that population. In linguistic experiments the participants taking part in the experiment and the test items under investigation are usually considered to be such random effects factors (cf. Baayen 2008:241). This is due to the following reasons. The participants of a production experiment may differ from each other in their ability to read out loud, their regional

background, their physical condition during the experiment or, when measuring pitch in particular, in their overall pitch range. All of these factors may more or less affect the speakers' performance and consequently the outcome of the experiment. Crucially, even if the surrounding experimental conditions are exactly the same for each speaker, they may still differ from each other with respect to these factors and many others, which are largely unknown to the researcher. Similarly, the constructed stimuli of such an experiment may differ from each other in sound structure, syllable length, the position of the highest prominent syllable in each constituent or the nucleus of the highest prominent syllable. The number of these external factors is indefinite and thus uncontrollable. Hence, both participants and items introduce a certain amount of variance to the data, which is unrelated to the experimental treatment factors of the experiment; not accounting for this between-speaker and between-item variability may obscure possible effects of the relevant treatment factors (cf. Kunter 2011). Furthermore, when replicating a given experiment, one would normally choose new speakers and items, which would again differ from the speakers and items of the first experiment in the above mentioned factors and various others, unknown to the researcher. Yet in a mixed-effects model, in which speaker and item can be modelled as random effects factors, the amount of variance introduced by these two variables is accounted for. Consequently, one can be more confident that significant effects in the fixed effects structure of the model are in fact associated with the experimental treatment factors.

In addition to that, mixed-effects models also have the advantage that they are robust with reference to unbalanced data sets (e.g. Pinheiro and Bates 2000). Unbalanced data sets occur when individual data points have to be excluded. If a statistical model fails to account for such unbalanced data sets, the statistical outcome is wrong. In the corpus study and the two experiments of the current thesis unbalanced data sets were likely to occur due to a failing of the automatic pitch tracking algorithm to return a pitch value for all measurement intervals. As noted earlier in section 3.2, one reason for a failing of the algorithm might be for instance a speaker's use of non-modal phonation such as creaky voice during one of the relevant measurement intervals. Furthermore, it might be possible that the length of a given measurement interval is too short for the algorithm to determine a pitch value for that particular segment (for more details on these two points the reader is referred to section 4.1.3). Consequently, the pitch algorithm might fail to return three pitch values for all compounds of the data set so that for some compounds we might only obtain two or just one value.

Besides, it is quite reasonable to believe that one also finds some outliers, i.e. ex-

treme data points, among the data. Such outliers may be caused for instance if the speakers used unexpectedly low or high pitches during the pronunciation of a compound. Yet, it might also be possible that the algorithm miscalculates the number of glottal pulses during a particular measurement interval, which results in either extremely low or high F0 values for that particular segment. It is recommended in the literature (e.g. Baayen 2008) to exclude such outliers during different stages of the modelling process in order to avoid them to obscure potential effects associated with the experimental treatment factors. The exclusion of such outliers might also lead to unbalanced data sets.

Because of the advantages of mixed-effects models over fixed-effects models, I decided to model the data of the current thesis by means of this particular type of model. The data was statistically analysed by means of the statistical software R (R Development Core Team 2011). The linear mixed-effects models were devised using the `lmer` function implemented in the statistical package `lme4` (Bates and Sarkar 2007). Details on the different fixed and random factors added to each model are given in the pertinent sections.

4 Corpus Study

This chapter presents a corpus study in which the prominence patterns of triconstituent NNN compounds are investigated by means of acoustic measurements of pitch.¹ The study investigates whether Liberman and Prince's (1977) Lexical Category Prominence Rule predicts the correct prominence pattern for left- and right-branching nominal compounds in English.

The chapter is structured as follows. Subsection 4.1.1 describes the corpus selected for the present analysis and provides details on the sampling procedure of the compounds. Subsection 4.1.2 deals with problems associated with the determination of the branching direction of the compounds, whereas subsection 4.1.3 describes the annotation process of the data and provides some more details on the pitch algorithm. Subsection 4.1.4 describes the predictions for the pitch analysis, i.e. how left- and right-branching compounds should differ from each other in pitch in order for the LCPR to be correct. The chapter continues with a subsection on the statistical modelling of the data before turning to the results of the analysis, which are presented in section 4.2. The results are further discussed in section 4.3. Finally, section 4.4 summarizes the findings and draws a first conclusion.

4.1 Methodology

4.1.1 The data: The Boston University Radio Speech Corpus

The data used in this study are taken from the Boston University Radio Speech Corpus (BURSC), an audio corpus collected by Ostendorf et al. (1996). The corpus consists of radio news texts from seven professional FM radio news speakers (4 male and

¹An earlier version of the corpus study described in this chapter has already been presented in Kösling and Plag (2009). The present version differs from Kösling and Plag's (2009) version with respect to its methodology. Kösling and Plag (2009) followed the method employed in Plag (2006) and calculated pitch differences between the three constituents of the compounds in order to account for prominence differences between left- and right-branching compounds. In this thesis the calculation of pitch differences is discarded. Instead, absolute pitch values are used, following the method employed in Kunter (2011).

3 female) all associated with the public radio station WBUR. The main portion of the corpus consists of more than seven hours of news recordings gained in the WBUR radio studio during actual broadcasts over a two-year period. In addition to the live recordings, the corpus also consists of a portion of 24 news stories (“lab news portion”) read by six of the seven speakers in a laboratory at the University of Boston. For these recordings, the speakers were first asked to read the news stories in their natural speech style and then, 30 minutes later, to read the same stories in their professional radio news style. Each story read by the news speakers has been digitized in paragraph size units, which typically include several sentences. All files are 16-bit recordings digitized at a 16 kHz sample rate (cf. Ostendorf et al. 1996).

The Boston University Radio Speech Corpus was chosen for this study because of the following reasons. First, the corpus was collected primarily to support text-to-speech synthesis, in particular the generation of prosodic patterns, and is thus ideally suited for the study of prosodic phenomena such as compound prominence. Second, the corpus contains data from the news genre, which I expected to contain a fair number of NNN compounds as they allow to capture complex information in a condensed form, which is particularly necessary in news texts. Third, it was assumed that professional news speakers tend to produce rather error-free speech. Finally, the corpus has already been used for research on prominence assignment of biconstituent NN compounds in studies by Kunter and Plag (2007), Lappe and Plag (2007), Plag et al. (2008), Plag and Kunter (2010), Plag (2010) and Kunter (2011) and proved to be highly suitable for this type of investigation. It was expected that the same would also hold for the investigation of triconstituent compounds.

The data for the present study were manually extracted from the text files of the corpus. In general all structures that formed a sequence of exactly three adjacent nouns within an NP were selected as potentially pertinent data. Some restrictions, however, were applied with reference to certain types of NNN sequences in order to ensure that the structures investigated conform as closely as possible to what most linguists would consider a triconstituent compound. Thus, NNN sequences containing initials such as *U.S. district judge* have been excluded from the analysis since the status of the abbreviation *U.S.* as a single noun-constituent is questionable.² In addition, I, rather conservatively, also excluded NNN structures that contained words other than English such as *Hillside hacienda*, *classroom blitz* or *San Antonio Spurs*. Yet, this restriction

²Thus it could be argued that *U.S.* is actually a compound itself, which would turn *U.S. district judge* into a four-constituent compound. Although I would not subscribe to such an analysis I wanted to restrict my analysis to items that are as uncontroversial as possible as regards their status.

did not apply to neoclassical formations, as for instance *biotechnology*, whose elements are of Greek or Latin origin. I decided to keep neoclassical formations in the data as their elements are deeply rooted in the English language and frequently used to coin new English words (e.g. Plag 2003). Yet, with reference to neoclassical formations like *biotechnology* a second question arose, namely whether these words should be themselves treated as compounds or rather derivatives. In the classical sense these formations are compounds and in terms of their semantic properties they behave like English compounds (cf. Plag 2003:74). Yet, in contrast to the components of English compounds, which are free morphemes, neoclassical elements are obligatorily bound, a property that they rather share with English affixes. Given that neoclassical formations share properties with both English affixes and compounds their status is ambiguous and the question arises as to whether a formation like *Massachusetts biotechnology council* should be treated as a triconstituent compound or not. I decided that formations like *biotechnology* should be included in the analysis as one constituent since such formations do not themselves consist of free elements as it is the case for English compounds. Hence, a formation like *Massachusetts biotechnology council* was treated as a triconstituent compound in the present thesis. Furthermore, NNN structures with genitive inflections, for example *tenant's right crisis*, were excluded from the analysis, and this policy was also applied to NNN sequences with proper names as the first two constituents like *Thomas Crown affair* and *John Hopkins University*. Moreover, compounds that contained constituents whose word class status was ambiguous were excluded from the data set. This was for instance the case for *cross* in *crossword puzzle*, *watch* in *watchdog group* or *expert* in *expert flavour analyst*. Finally, it was also necessary to decide whether words that contained *man* as their second element, such as *congressman*, *chairman* or *postman*, should be treated as an NN compound or as a derivative, as it has been claimed by some authors (e.g. Marchand 1969) that in such words *man* has lost its status of a free lexical morpheme and instead functions as a suffix. The suffix-like status of *man* in these words is argued to be indicated by the fact that in *man* the vowel is usually reduced to a schwa. Such vowel reduction is often characteristic of a change from a free morpheme to a bound morpheme (cf. Hoeksema 1985:68).³ Yet, in this thesis, *man* in words like *chairman* or *congressman* was not treated as a suffix but as a free lexical morpheme and thus compounds such as *state chairman* or *Massachusetts congressman* would enter the analysis as a triconstituent compound.

The sampling procedure of the compounds was as follows. Starting with the tran-

³Marchand (1969:356) refers to *man* as being a semi-suffix, which is an element that stands midway between a lexical morpheme and a suffix.

scription of speaker F1, all NNN compounds that conformed to the above mentioned restrictions were extracted. For each type, only the first token was sampled such that additional tokens did not enter the analysis.⁴ Furthermore, plural and singular forms of one type were treated as one type. The data from the other speakers were sampled in the following sequence: F2, F3, M1, M2, M3, M4. Table 4.1 gives the distribution of the types sampled across speakers.⁵

Table 4.1: Number of types sampled across speakers

Speaker	Number of Types
F1	69
F2	57
F3	123
M1	57
M2	84
M3	20
M4	94

Applying the procedures just described, we ended up with a data set of 505 NNN structures.

4.1.2 Branching direction

In order to test the predictions of the LCPR, the 505 NNN structures extracted from the corpus had to be coded according to their internal structure, i.e. as either left- or right-branching. Crucially, the analysis of the branching direction was performed on the basis of the written transcript alone. Listening to the news stories was avoided in order not to confound the analysis of the branching direction with acoustic information on prominence. How did I determine the branching direction of a given tri-constituent compound? I performed a semantic analysis of all 505 compounds. The semantic analysis of the majority of these compounds was rather straightforward and

⁴It is well-known, though not well researched (but see Bauer 1983a:103; Kunter 2011), that there is sometimes variation in prominence across tokens of the same compound. Taking just the first instance of each compound therefore runs the risk of losing interesting data, as well as losing an opportunity to assess this type of variation. However, I wanted to test the LCPR under its own assumptions, in particular under the assumption that we abstract away from within-type variability. Taking only one token per type has the additional advantage that many variant tokens of a limited number of compounds do not unduly influence the overall distributions. A larger study is certainly called for that tests within-type variability of triconstituent compounds.

⁵The labels 'F' and 'M' mark the speakers' sex, with 'F' referring to female speakers and 'M' indicating that the data come from a male speaker.

led to a set of 448 either clearly left-branching, or clearly right-branching compounds, such as [*seat belt*] *law* or *state* [*income tax*].

For 85 compounds among the 448 compounds an additional independent criterion was available to determine the branching direction, i.e. the orthographic representation. In these compounds two of the constituents were either written as one orthographic word or with a hyphen. Although the spelling of compounds varies among speakers, it is clear that a more intricate spelling, i.e. as one word or hyphenated, is an indication of a more word-like status of that combination.⁶ Hence, I would expect that a more intricate spelling indicates the presence of a complex constituent, as illustrated in, for example, *weekend series*, *wheelchair marathoners*, *Boston newspaper* and *company whistle-blowers*. I used this insight to verify my semantic analysis in the following way. I assigned a branching direction to this subset of compounds based on their spelling, and I then checked the results of the application of the spelling criterion against the results of the application of the semantic criterion. This resulted in a 100% match between the two criteria.

The semantic analysis of the extracted compounds, however, did not always yield such clear-cut interpretations as the ones just mentioned. It turned out to be rather problematic for structurally ambiguous compounds. Structurally ambiguous compounds such as *silver knife handle* may be interpreted as left-branching ('the handle of a silver knife'), or as right-branching ('knife handle made of silver').

It is usually assumed that such ambiguity arises primarily when compounds occur in isolation. As soon as they are embedded in a natural speech context, one can usually interpret them unambiguously with reference to that context (see, for example, Meyer 1993 and Plag 2003 for some discussion). It was for this reason that the number of truly ambiguous compounds was expected to be extremely small at first, since all compounds used in this study were embedded in a natural speech context. Nevertheless, it turned out that for 57 compounds of the 505 compounds even the context could not provide enough information to clearly disambiguate them. For instance, a *Boston police officer* may be an officer of the Boston police (left-branching), or it may be a police officer working in Boston (right-branching).⁷

These ambiguous compounds were excluded from the analysis, which reduced the number of items under investigation to 448 compounds. Of the 448 NNN compounds, 326 were classified as being left-branching, 122 as right-branching. The high propor-

⁶See Plag et al. (2007), Plag et al. (2008) and Sepp (2006) for more discussion and evidence.

⁷Although the meaning difference between two interpretations may in fact be rather subtle, the LCPR would nevertheless predict different prominence patterns for the two differently branching structures.

tion of left-branching compounds in contrast to right-branching compounds is not peculiar to my data. The result goes into the same direction as earlier findings by Marchand (1969), Warren (1978) or, more recently, Berg (2006). Based on their findings these authors claim that left-branching compounds are more common than right-branching compounds in English, with left-branching compounds being the unmarked structure for triconstituent compounds.

4.1.3 Acoustic measurements

The 448 left- and right-branching compounds extracted from the corpus were annotated using the speech analysis software PRAAT (Boersma and Weenik 2007). Following the method employed by Plag et al. (2008), in PRAAT textgrids, I first manually segmented the single constituents of each compound and second the sonorant part of the rime of the most prominent syllable in each of the three compound constituents. The segmentation was done on the basis of a combined wave and spectrogram view. The segment boundaries were determined based on the criteria provided for consonant and vowel segmentation in Ladefoged (2003). Furthermore, during the segmentation of the most prominent syllable in each constituent, special attention was paid to potential occurrences of prominence shifts. A shift of prominence may occur when a lexical item is embedded in a more complex structure and the item's most prominent syllable immediately precedes the most prominent syllable of the following constituent. In this case the prominence of the first constituent may be shifted to the left in order to avoid a so-called clash with the following prominent syllable (e.g. Liberman and Prince 1977; Giegerich 1985). To give an example, the item *Massachussets* carries highest prominence on the third syllable when spoken in isolation, yet when embedded in a more complex structure such as *Massachussets miracle* prominence may be shifted to the first syllable of the word (*Mássachussetts miracle*) (cf. Shattuck-Hufnagel et al. 1994:359). In the present data set, such prominence shifts were for instance noted in the compounds *Mássachussets school children* or *Mássachussetts wage earners* and were segmented accordingly.

The mean F0 value of the selected interval for each constituent was automatically measured with the help of a PRAAT-script. As noted earlier in section 3.1, I used the same PRAAT-script as Kunter (2011) in his study. The script took the standard values proposed for a pitch analysis by the PRAAT programme as a baseline. Hence, gender-specific pitch ranges were considered by choosing pitch boundary settings of 75-300

Hz for male speakers and 100-500 Hz for female speakers.⁸ Automatic adjustments were made in cases of creaky voice or octave jumps, in case a pitch contour could be extracted only for half or less of an interval as well as if the minimal pitch extracted from a given interval was less than 0.5 semitones higher than the pitch floor setting (i.e. 75Hz for males and 100Hz for females). In cases in which adjustment was necessary due to the factors just mentioned, the settings for pitch floor and ceiling values were reduced automatically by one third, and the voicing threshold was reduced to increase the sensibility of the pitch extraction algorithm. The pitch measurement was repeated with the new settings up to three times. If after three adjustments it was still impossible to detect a pitch value, the affected observations were excluded from the analysis (for more details regarding the algorithm applied see Kunter (2011:103f)). For the 448 triconstituent compounds extracted from the corpus and classified according to branching-direction, the script was supposed to measure three pitch values, i.e. one for each constituent. Yet, for six compounds, the algorithm only returned two pitch values so that the total number of pitch values measured by the script was 1338 rather than 1344.

In a subsequent procedure, the 1338 pitch values were logarithmically transformed from Hertz (Hz) into semitones (ST) relative to the lowest pitch value measured in the data set (cf. Kunter 2011). The transformation of the pitch values into semitones was done by means of the formula given in 4.1. In the formula f_i is the i -th mean pitch measurement, and $minf$ the minimum of all observed mean pitch measurements.

$$(4.1) \quad f_{i,ST} = \frac{\log(f_i/minf)}{\log(2)}$$

Having measured the mean pitches in each of the three members of left- and right-branching compounds, the question that arises next is that of how left- and right-branching compounds should differ from each other in their pitch pattern. I will deal with this question in the following subsection.

4.1.4 Prediction: Lexical Category Prominence Rule

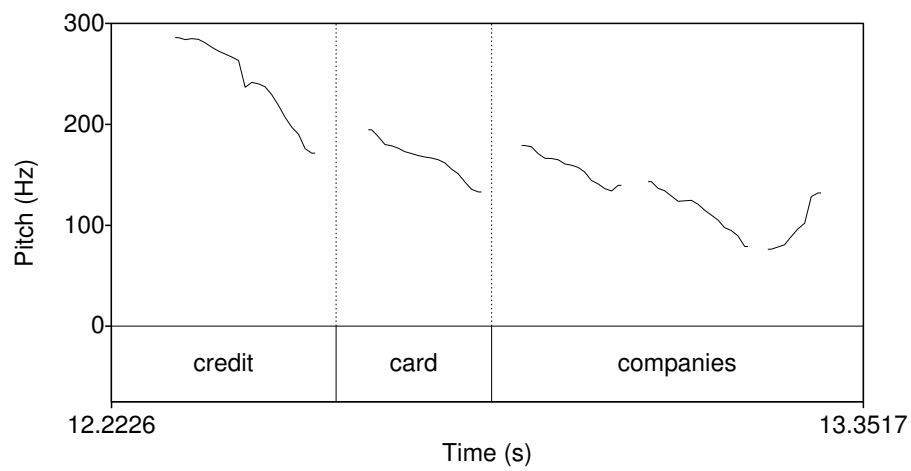
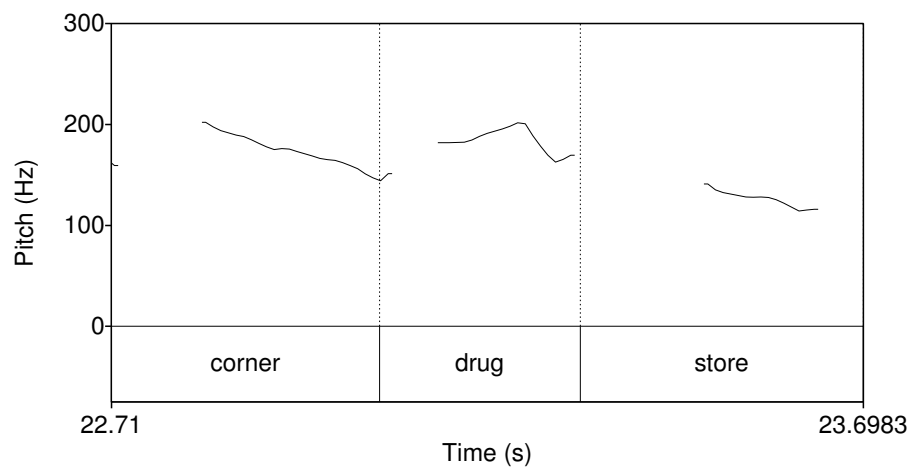
Testing the LCPR by means of pitch analysis raises the question of how we know that left-branching compounds exhibit highest prominence on constituent N1 and right-branching compounds on constituent N2. Which pitch patterns do we expect for left-

⁸75Hz and 100Hz, respectively, represent the pitch floor, i.e. the bottom of the pitch range selected for male and female speakers. The pitch ceiling refers to the top value of the selected pitch range, i.e. 300Hz for males and 500Hz for females, respectively.

and right-branching compounds? In order to understand the predictions for triconstituent compounds, we should briefly repeat what was noted earlier about pitch and prominence for biconstituent NN compounds. As noted in section 2.1, studies by Farnetani et al. (1988), Plag (2006), Plag et al. (2008) and Kunter (2011) dealing with prominence assignment in biconstituent compounds have found that left prominent NN compounds are characterized by a high pitch on constituent N1 and a clearly lower pitch on constituent N2. In contrast, right prominent NN compounds were found to have a relatively high pitch on both constituent N1 and constituent N2. Crucially, the studies also showed that pitch on the right member of a right prominent NN compound is not necessarily higher than on the left constituent. Instead, it was found that in right prominent NN compounds pitch on the right constituent was in general slightly lower than on the left constituent. The reason for this slight pitch drop between the left and right member of a right prominent compound was argued to be the result of a general pitch declination effect (cf. Kunter 2011). Yet, according to Gussenhoven (2004) and Kunter (2011) listeners are able to make up for this pitch declination effect and perceive the right member of a right prominent compound as more prominent than the left member despite its lower pitch. Given these findings for biconstituent NN compounds, we may now turn to the predictions for triconstituent compounds.

The LCPR predicts that in left-branching compounds highest prominence is assigned to constituent N1 whereas in right-branching compounds highest prominence is assigned to constituent N2. According to this prediction, we would expect the following two pitch patterns for left- and right-branching compounds. Left-branching compounds should exhibit a clearly higher pitch on constituent N1 than on constituent N2 and constituent N3 since constituent N1 should be the most prominent constituent in these compounds. In contrast, in right-branching compounds the prominence relation between constituent N1 and constituent N2 is that of right prominence as constituent N2 is expected to be more prominent than constituent N1. Hence, for right-branching compounds we would expect a high pitch on both constituent N1 and constituent N2. The lowest pitch should be assigned to constituent N3 in right-branching compounds, as constituent N3 is expected to be the least prominent constituent.

For a better illustration of the predicted pitch behaviour of left- and right-branching compounds, Figures 4.1 and 4.2 show two pitch tracks. In particular, Figure 4.1 shows the pitch track of the left-branching compound [*credit card*] *company* and Figure 4.2 that of the right-branching compound *corner* [*drug store*]. The two compounds are se-

Figure 4.1: Pitch track of a left-branching compound: [*crédit card*] *companies*Figure 4.2: Pitch track of a right-branching compound: *corner* [drúg store]

lected from the corpus data. They exhibit a pitch pattern as expected under the LCPR. The vertical axis gives the pitch range in Hertz (Hz), the horizontal axis displays the duration of the utterance. The line in the middle of the plot shows the pitch curve in each compound as measured by the PRAAT pitch tracking algorithm. The pitch track of the left-branching compound [*credit card*] *company* shows that highest pitch is assigned to constituent N1, which clearly decreases over the utterance of constituent N2 and constituent N3. In contrast, in Figure 4.2, we observe a relatively high pitch on both constituent N1 and constituent N2. The high pitch on constituent N2 is due to a small increase in pitch on that constituent. With reference to the pitch assigned to constituent N3 in the compound *corner* [*drug store*], we observe that it is clearly lower than that on the first two constituents of the compound. As it becomes clear from the two pitch tracks, left- and right-branching compounds are expected to differ in pitch assigned to constituent N2. Hence, a comparison of the pitch values of left- and right-branching compounds should reveal a higher pitch on constituent N2 in right-branching compounds than in left-branching compounds. In contrast, with reference to the pitches assigned to constituent N1 and constituent N3 we would not necessarily expect left- and right-branching compounds to differ from each other. Yet, it may be possible that pitch on constituent N1 in right-branching compounds will be slightly lower than in left-branching compounds. In his study on the acoustic correlates of compound stress, Kunter (2011) found that apart from a higher pitch on the right member, right prominent compounds also tended to have slightly lower pitches on constituent N1 than left prominent NN compounds. Given this result, it may be possible that in the present study right-branching compounds also tend to have slightly lower pitches on constituent N1 than left-branching compounds instead of having equally high pitches. Yet, whether right-branching compounds have lower or equally high pitches on constituent N1 than left-branching compounds is less crucial with reference to the LCPR prediction. In either case, right-branching compounds should exhibit a higher pitch on constituent N2 than left-branching compounds.

4.1.5 Statistical procedure

Before I fitted a mixed-effects model to the data, I inspected the pitch distribution of each individual speaker in order to see whether there were any extreme data points, i.e. outliers, among the pitch values. The inspection of the pitch distribution of each speaker was in so far important as such outliers may affect initial model fitting stages if not accounted for (cf. Baayen 2008:244). In the present corpus study outliers might

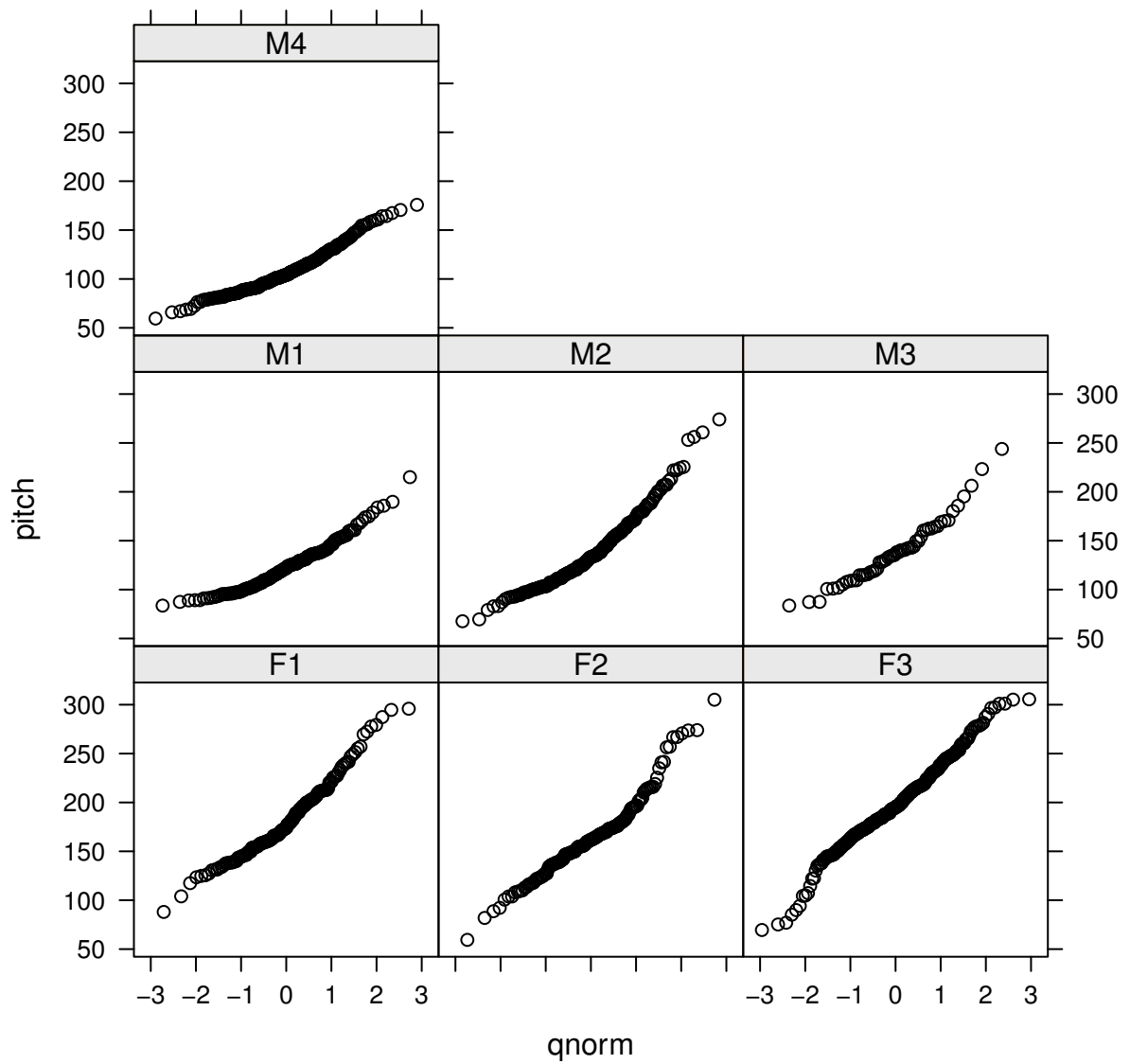
have been due to technical reasons such as measurement errors by the script. In addition, when measuring pitch in particular, one runs the danger of pitch measurements being affected by the speakers' use of creaky voice during the utterance of the compounds. In creaky voice the vocal cords are pressed together more tightly than in modal phonation, which leads to irregular vibrations of the vocal cords and thus to a lower rate of glottal pulses within a speech segment (Ladefoged 2003:chapter 7). As a result, the fundamental frequency of the speech signal is extremely lowered in segments affected by creaky voice and these low pitch values may turn up as outliers in the data set.

The closer inspection of the pitch distributions for each speaker of the corpus indeed revealed a few strikingly low and high pitch values for some of the speakers. This is illustrated in Figure 4.3. The Figure shows a quantile-quantile plot that displays the pitch distribution for each individual speaker. The plot shows for instance that speaker M2 has four values which are notably higher than the rest of his values. Furthermore, for speaker F3, we observe that the left tail of her distribution forms a small curve instead of a straight line. The inspection of the lower pitch values for each speaker revealed that these values were in fact mainly instances of creaky voice phonation. Since the present thesis concentrates on the F0 (pitch) as the only acoustic correlate of compound prominence, such unnaturally low F0 values caused by creaky voice are highly problematic, as they directly affect the target value. In particular, it is plausible that leaving creaky voice values in the data may lower the average mean F0 of a given constituent to such an extent that it differs significantly in its mean F0 from another constituent. Yet, the difference would not be the result of a prominence difference between two constituents but rather an artefact of different phonation modes. Because of this problem associated with creaky voice phonation, I decided to exclude those observations a priori to the statistical modelling of the data.⁹ This reduced the total number of pitch values calculated by the pitch algorithm from 1338 to 1307.

Besides, the closer inspection of the few relatively high values for some speakers (see again Figure 4.3) revealed that these observations were not caused by non-modal phonation or measurement errors. Because there was no technical reason to exclude these values, I decided to leave them in the data set at that point of the analysis in order to avoid unnecessary data trimming. Yet, I had good reasons to believe that if these high pitch values remained outliers in the later stages of model fitting, they

⁹It should be additionally mentioned that Kunter (2011) observed that creaky voice occurs more frequently in non-prominent constituents than in prominent constituents. Based on his result he concludes that creaky voice is not a marker of prominence.

Figure 4.3: Quantile-quantile plots for the pitch distribution grouped by speakers of the BURSC.



would be eliminated during the process of model criticism.¹⁰

I devised a linear mixed-effects model with *SPEAKER* and *ITEM* as random effects and the three categorical factors *BRANCHING*, *POSITION* and *GENDER* as fixed effects. The dependent variable in the model was that of pitch in semitones (*pitchST*). The random effect *SPEAKER* was included in the model in order to account for between-speaker differences, as for instance, differences in the pitch range between the seven speakers. Furthermore, I included *ITEM* as a second random effect in the model in order to account for the fact that the items were pronounced by male and female speakers. The inclusion of the two random effects was verified by means of two likelihood ratio tests (cf. Baayen 2008:253). In such likelihood ratio tests the model containing the random effect is compared to a simpler model without the respective factor fitted as a random effect to the model. If the test reveals a significant increase in the log likelihood for the model that contains the random effect it indicates that the inclusion of the random effect in that model is indeed justified. In the current model this was the case for both *SPEAKER* and *ITEM* so that their inclusion was verified (see Appendix part B for the two likelihood ratio tests).

As mentioned above, in addition to the two random effects, I fitted the three fixed factors *BRANCHING*, *POSITION* and *GENDER* to the model. The factor *BRANCHING* had the two factor levels *left* and *right*. The factor functioned as a grouping factor in the model and coded the information as to whether the compounds were left- or right-branching. The second factor *POSITION* consisted of the three factor levels *lpitch*, *mpitch* and *rpitch*. The three factor levels code the position in the compound for which pitch was measured. Thus, *lpitch* marks the pitch value obtained for the leftmost constituent of a compound, the labels *mpitch* and *rpitch* refer to the pitch values measured for the constituent in the middle (N2) and the rightmost constituent (N3), respectively. Finally, the third factor *GENDER* consisted of the two factor levels *female* and *male*. The factor was added to the model in order to account for the general difference in the pitch range between female and male speakers. As mentioned earlier in section 4.1.3, the pitch range of adult male speakers generally ranges between 75 and 300 Hz, whereas the pitch range for adult female speakers ranges between 100 and 500 Hz. Given these general pitch differences between the two genders, I expected this factor to be highly significant.

Finally, in addition to these three main effects, I added an interaction of *BRANCH-*

¹⁰In a study by Baayen and Milin (2010) dealing with the statistical modelling of reaction times, it is shown that a combination of a priori data trimming and additional model criticism increases the fit of the final model as opposed to approaches that only apply one of the two methods.

ING by POSITION. The inclusion of the interaction was crucial in order to find out whether left- and right-branching compounds differ in the degree to which pitch changes between the three constituents. Under the assumption that left-branching and right-branching compounds differ in pitch on constituent N2, we would expect a statistically significant difference in the pitch change between constituent N1 and constituent N2 of left- and right-branching compounds. This should be indicated by the presence of a statistically significant interaction of BRANCHING by POSITION. In the absence of a significant interaction of the two factors BRANCHING and POSITION, the model would not provide any evidence that left- and right-branching compounds differ from each other as predicted by the LCPR.

I fitted the above specified model to the data and subsequently checked the residuals of that model for potential outliers. Following the procedure employed for instance in Baayen (2008) and Baayen and Milin (2010), all data points with residuals greater a standard deviation than $|2.5|$ were excluded from the data set. This resulted in the exclusion of 17 data points (1.3 %), which reduced the number of pitch values from 1307 to 1290.¹¹ A new model was fitted to the trimmed data set. An inspection of the residuals of that new model showed that the non-normality of the residuals of the first model was reduced.

As for the random effects of the model, I also checked whether the inclusion of all the fixed effects in the model was justified. In order to do so, I compared the model with the two-way interaction of BRANCHING by POSITION with a simpler model not containing this interaction parameter (e.g. Crawley 2005b:chapter 7). The comparison of the two models revealed that the interaction parameter had to be kept in the model. The same procedure was repeated for the factor GENDER, which also emerged as significant. The significance of GENDER indicates that - as expected - male and female speakers differ from each other with reference to their general pitch height. Furthermore, the significant interaction of BRANCHING by POSITION suggests that left- and right-branching compounds do indeed differ in the degree to which pitch changes between the three constituents. Yet, in order to know at which position, i.e. between which constituents pitch changes to a different degree in left- and right-branching compounds, we have to look at the table of coefficients of the model as presented in the next section.

¹¹The exclusion of the 17 outliers resulted in a loss of the left-branching compound *rental housing association*. For this compound no pitch value remained in the final data set after the a priori data trimming and the final model criticism. Hence, 447 compounds entered the final analysis, i.e. 122 right-branching compounds and 325 left-branching compounds. For 414 of the 447 compounds three pitch values were available, with 33 compounds having only one or two values.

4.2 Results

4.2.1 Testing the LCPR

The central aim of the present analysis is to test to what extent the predictions made by the Lexical Category Prominence Rule really hold for English NNN compounds. For this reason I measured the mean pitch over the sonorant part of the rime of the most prominent syllable in each of the three constituents of left- and right-branching compounds and compared these values by means of a linear mixed-effects model. In this section the results of the statistical analysis are presented. The p-values provided in this section were obtained by means of Markov chain Monte Carlo sampling (MCMC) (cf. Baayen 2008:248).

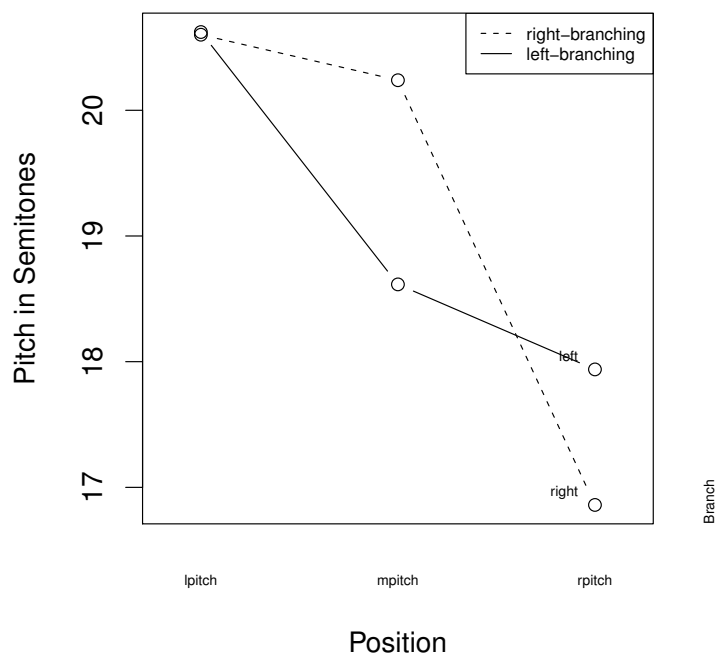
Table 4.2 shows the table of coefficients, i.e. the fixed effects coefficients, of the final model (for the full model, i.e. the model that also includes the random effects structure, see Appendix part B). The baseline in table 4.2, i.e. the mean value that is mapped on the model's intercept, is that of the average mean pitch of constituent N2 (*mpitch*) in left-branching compounds (*left*) obtained for the female portion of the data set (*female*). Figure 4.4 shows the respective interaction plot of the model. Pitch in semitones (*pitchST*) is given on the vertical axis of the plot; the three positions, i.e. the three constituents for which pitch was measured, are displayed on the horizontal axis. The dashed line connects the three mean values of the right-branching compounds, the solid line those obtained for the left-branching compounds. The mean values are marked by the circles in the plot. The plot shows the three mean values of left- and right-branching compounds obtained for the female portion of the data set. The corresponding pitch values for the male speakers of the data set are in general about 6 semitones lower for both left- and right-branching compounds. This information is provided by the coefficient of *genderm* in table 4.2. The negative coefficient tells us that we need to subtract about 6.4 semitones from the baseline of the model in order to obtain the male value for constituent N2 in left-branching compounds. Given that *GENDER* is fitted as a main effect to the model, the same value must be generally subtracted from the female values in order to obtain the respective pitch values for the male speakers.

The interaction plot of the model indicates that both left- and right-branching compounds start off with roughly the same pitch (20.6 ST) on constituent N1. However, with reference to the relevant mean pitches of constituent N2, we observe a clear difference between left- and right-branching compounds. In particular, as predicted by the LCPR, we observe a higher pitch on constituent N2 in right-branching compounds

Table 4.2: Table of coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: BRANCHING = left, POSITION = mpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	18.6149	1.1094	16.779	< 0.001
right	1.6244	0.3299	4.925	< 0.001
lpitch	2.0078	0.2122	9.463	< 0.001
rpitch	-0.6765	0.2148	-3.149	< 0.01
male	-6.3911	1.4596	-4.379	< 0.001
right:lpitch	-1.6456	0.4034	-4.079	< 0.001
right:rpitch	-2.7026	0.4094	-6.601	< 0.001

Figure 4.4: Interaction plot displaying the average mean pitches estimated by the model for each constituent of left- and right-branching compounds; female speakers.



(20.3 ST) than in left-branching compounds (18.6 ST). This higher pitch on constituent N2 in right-branching compounds is due to a relatively small pitch drop of just 0.3 semitones between constituent N1 and constituent N2 in these compounds as opposed to a pitch drop of 2 semitones between constituents N1 and N2 in left-branching compounds. In addition to that, we observe that the lowest pitch is generally assigned to constituent N3 in the two groups. Yet, we also note that left-branching compounds tend to have a slightly higher pitch on constituent N3 than right-branching compounds due to a smaller overall pitch range in left-branching compounds. In particular, pitch drops on average about 3.2 semitones over the entire utterance of right-branching compounds but only 2.68 semitones in left-branching compounds.

The observation that left- and right-branching compounds start off with an equally high pitch, but that right-branching compounds exhibit a higher pitch on constituent N2 than left-branching compounds is in accordance with the LCPR prediction. A closer look at Table 4.2 shows that this visually gained impression is also statistically supported. Table 4.2 lists all differences between the pitch mean mapped on the intercept of the model (`mpitch`, `female`, `left`) and the means of the other factor levels of the three factors `BRANCHING`, `POSITION` and `GENDER`. Differences between factor levels that do not involve the model's baseline are not displayed in the table; yet they may be obtained by altering the order of the factor levels by means of treatment contrasts.¹² Due to the selected baseline, Table 4.2 lists the coefficient that provides the difference between the average pitch means obtained for constituent N2 in left- and right-branching compounds `right`. Furthermore, by adding the coefficients of `lpitch` and `rpitch` to the intercept, we obtain the mean pitch of constituent N1 and the mean pitch of constituent N3 of left-branching compounds, respectively. Apart from that Table 4.2 also lists all interactions that involve the group mean mapped on the intercept of the model. Thus, the two interaction coefficients listed in table 4.2 tell us to what extent left- and right-branching compounds differ in the pitch change between constituent N1 and constituent N2 on the one hand (`right:lpitch`) and between constituent N2 and constituent N3 on the other hand (`right:rpitch`). The interaction coefficient relevant for the present analysis is thus the one of `right:lpitch`. The interaction coefficient indicates that the pitch change between constituent N1 and

¹²The R-programme uses treatment contrasts as its default convention. This means that the factor level that comes alphabetically and numerically first is automatically chosen as the model's baseline. In the present data set the automatically selected baseline was that of `lpitch`, `left` and `female`. If necessary, the baseline of the model may be changed by changing the treatment contrasts, i.e. by altering the order of the factor levels of `BRANCHING`, `POSITION` and `GENDER`, respectively. This was done in Table 4.2 in order to directly display the relevant comparison between the N2 pitches of left- and right-branching compounds.

constituent N2 in right-branching compounds is indeed significantly smaller than in left-branching compounds ($p < 0.001$). Furthermore, the coefficient of *right* tells us that pitch on constituent N2 in right-branching compounds is 1.6 semitones higher than in left-branching compounds and that this difference is also statistically significant ($p < 0.001$).¹³ The result is in accordance with the LCPR prediction.

Finally, we note that apart from the predicted difference between the N2 pitches between left- and right-branching compounds, the model also reveals a statistically significantly higher pitch on constituent N3 in left-branching compounds than in right-branching compounds ($p < 0.01$).¹⁴ This higher pitch on constituent N3 in left-branching compounds is rather unexpected as we would have assumed that left- and right-branching compounds only differ in pitch assigned to constituent N2. Thus, we may raise the question of why we find this general tendency for left-branching compounds to have a higher pitch on constituent N3 than right-branching compounds. A plausible explanation for this difference might be that some of the left-branching compounds have highest prominence on constituent N3, which is marked by a relatively high pitch on that constituent. As argued in section 2.1, the presence of such compounds in the corpus would be relatively unsurprising, given that left-branching NNNs with highest prominence assigned to constituent N3 are attested in the literature (e.g. Fudge 1984; Liberman and Prince 1977; Liberman and Sproat 1992; Hayes 1995; Sproat 1994), although they are referred to as being phrases by most authors.

To sum up, for right-branching compounds the analysis revealed a high pitch on both constituent N1 and constituent N2 and a clearly lower pitch on constituent N3. In contrast, for left-branching compounds, we found a high pitch on constituent N1 but clearly lower pitches on both constituent N2 and constituent N3. As predicted, the pitch comparison of the two compound groups revealed a generally higher pitch on constituent N2 in right-branching compounds than in left-branching compounds. Hence, the analysis provides evidence that speakers systematically assign higher pitches, i.e. higher prominence, to constituent N2 in right-branching compounds than in left-branching compounds, a result fully in accordance with the LCPR prediction.

Finally, with reference to our result, however, it is also important to note again that - based on the present analysis - we cannot make absolute statements in the sense of

¹³An additional inspection of the group means of constituent N1 of left- and right-branching compounds provided no evidence for the two compound groups to differ systematically from each other in pitch assigned to that constituent (see Appendix part B for the table of coefficients displaying the contrast between the N1 pitches of left- and right-branching compounds.)

¹⁴The table of coefficients that displays the difference between the mean pitches of constituent N3 between left and right-branching compounds is given in the Appendix part B.

that all left- and right-branching compounds in the corpus actually behave according to the LCPR. The reason for this is that gradient measures, i.e. statistically significant differences between left- and right-branching compounds, only enable us to determine general tendencies in a certain direction. Thus, despite this general trend of left- and right-branching compounds to behave as predicted by the LCPR, we cannot entirely rule out that there may also be some compounds in the data set that do not behave according to the LCPR. In fact, given the findings of previous studies by Kvam (1990), Sproat (1994), Berg (2009), Giegerich (2009) on the one hand, and the theoretically problematic assumption on which the LCPR is based, we have good reasons to believe that we may also find some violations among the data. In fact, a first hint towards such violational patterns to the LCPR may be seen in the general tendency for left-branching compounds to have a higher pitch on constituent N3 than right-branching compounds. In view of this situation, I decided to explore the corpus data in more detail in order to find out about potential violators of the LCPR. The exploration of the data is described in the following subsection.

4.2.2 Variation and branching direction

In this subsection I want to address the question of how many items within the two groups might violate the LCPR, and what kinds of violations one may find (at the IC-level, or at the N-level). Using gradient measures, however, makes this a difficult task as the only way to make out clear subgroups of compounds that violate the LCPR, would be to classify all compounds as either being prominent on constituent N1, constituent N2 or constituent N3. Hence, the crucial question arises of how I know that in a given compound constituent N1, constituent N2 or constituent N3 is the most prominent constituent.

One way to determine aberrant cases to the LCPR would be that of measuring pitch differences between each of the three members of left- and right-branching compounds (e.g. Farnetani et al. 1988; Plag 2006). By calculating pitch differences, we obtain a direct measure for the pitch relation between each of the three constituents of the triconstituent compounds. Note that it was found that left prominent NN compounds have a relatively high pitch on constituent N1 and a clearly lower pitch on constituent N2. In contrast, right prominent NN compounds were found to have a relatively high pitch on both the left and right constituent. Hence, left prominence should be indicated by a large pitch difference and right prominence by a smaller pitch difference. However, calculating pitch differences alone does not solve the problem of determin-

ing exceptions to the LCPR, as we would still need to define a suitable threshold for these pitch differences along which we could clearly separate cases of violations from cases that conform to the rule. This raises the question of which threshold is appropriate? At which point can we more or less clearly say that the relation between two constituents reveals left prominence or right prominence, respectively?

As argued in section 4.1.4, one can generally assume that whenever a constituent to the right has a higher pitch than a constituent to the left, the constituent to the right is more prominent. Thus, we can assume that negative pitch differences clearly indicate right prominence between two constituents. But what would be a suitable threshold for left prominence? As noted before, right prominence is not necessarily indicated by a higher pitch on the right constituent, but right prominence may also be indicated by a slightly lower pitch on the right constituent due to a general pitch declination effect (e.g. Kunter 2011). Because of that, right prominence may also be indicated by a small positive pitch difference between two elements in a compound (cf. Plag 2006). Given this fact, a threshold of 0 ST for left prominence, would not be suitable as we would run the danger of classifying compounds that present exceptions to the LCPR as actually conforming to the rule or vice versa. Yet, under the assumption that the more positive the pitch difference, the clearer the left prominence, we should choose a pitch difference that is not too close to 0 ST. A look at the means in table 4.2 shows us that the difference between the mean of constituent N1 and constituent N2 in left-branching compounds is 2.06 ST. Recall that according to the assumptions of the LCPR the pitch difference between constituent N1 and constituent N2 should reflect left prominence. I therefore decided to take 2 ST as the threshold for left prominence, i.e. I assume that values equal or above 2 ST indicate left prominence. This entails that I choose not to say anything about the prominence relationship for all those constituent pairs whose pitch difference is between 0 and 2 ST. While I am losing data under this approach, I try to minimize the risk of making wrong generalizations. In other words, by using this methodology, I try to be conservative and rather underestimate the number of exceptions to the LCPR.

The pitch differences were calculated by means of the formula given in 4.2, which at the same time logarithmically transformed the pitch differences from Hertz into semitones (Δ_{ST}) (cf. e.g. Henton 1989:302). In the equation, f_x refers to the left one of two pitches, whereas f_y refers to the right one. Hence, when calculating the pitch difference between constituent N1 and constituent N2, f_x functions as a placeholder for the pitch value of constituent N1, and f_y refers to the pitch value obtained for constituent N2. Accordingly, when calculating the pitch difference between constituent

N2 and constituent N3, f_x stands for the pitch value of constituent N2 and f_y for that of constituent N3.

$$(4.2) \quad \Delta_{ST} = 12 * \log(f_x/f_y)/\log 2$$

The transformation into semitones was necessary in order to neutralize gender-specific pitch differences. The three pitch differences calculated for each compound were labelled 'P1P2', 'P2P3' and 'P1P3'. P1P2 designates the pitch difference between constituents N1 and N2, P2P3 the difference between constituents N2 and N3, and P1P3 the one between constituents N1 and N3. As the calculation of these pitch differences required that for each compound three pitch values were available, i.e. one pitch value for each of the three compound constituents, only 414 of the 447 compounds qualified for a closer inspection (301 = left-branching compounds; 113 = right-branching compounds).¹⁵ Furthermore, of these 414 compounds, 98 left-branching and 39 right-branching compounds turned out to have at least one pitch difference between 0 and 2 ST. As mentioned in the previous paragraph, I abstained from making any statement about the prominence pattern of such compounds.

I first turn to the prominence patterns I found for left-branching compounds. According to the LCPR highest prominence in left-branching compounds is assigned to constituent N1. Thus, the LCPR is violated as soon as highest prominence is assigned to constituent N2 or constituent N3. Highest prominence on N2 violates the LCPR at the N-level but not at the IC-level since in that case highest prominence remains within the complex constituent, with node B being weak and node A being strong. Highest prominence on N3 violates the LCPR at the IC-level, since it causes node B to be strong, in spite of its being non-branching. The patterns that violate the LCPR are listed in Table 4.3, in which POSITIVE means $ST > 2$ (indicating left prominence), and NEGATIVE means $ST < 0$ (indicating right prominence). The label IRRELEVANT indicates that the respective pitch difference is not relevant in order to determine the most prominent constituent in these compounds; hence the pitch difference may be either positive or negative.

Pattern 1 indicates IC-level violations for left-branching compounds. In pattern 1 the negative P2P3 difference indicates that N3 is more prominent than N2. Furthermore, the negative P1P3 difference indicates that N3 is also more prominent than constituent N1. Hence, constituent N3 is the most prominent constituent in these compounds. As shown in table 4.3 the pitch relation between constituent N1 and

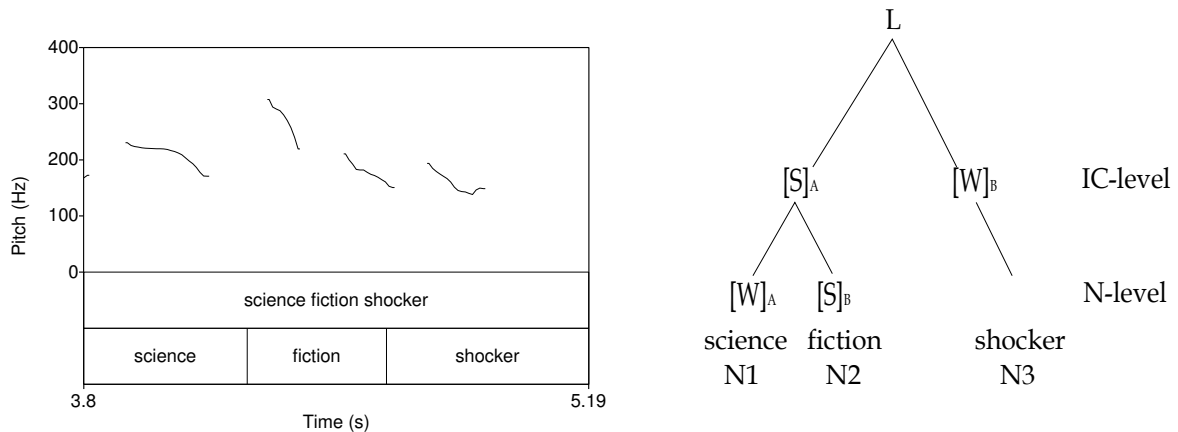
¹⁵As discussed in subsections 4.1.3 and 4.1.5 some pitch values had to be excluded, for example, due to creaky voice.

Table 4.3: Violations of the LCPR: left-branching compounds

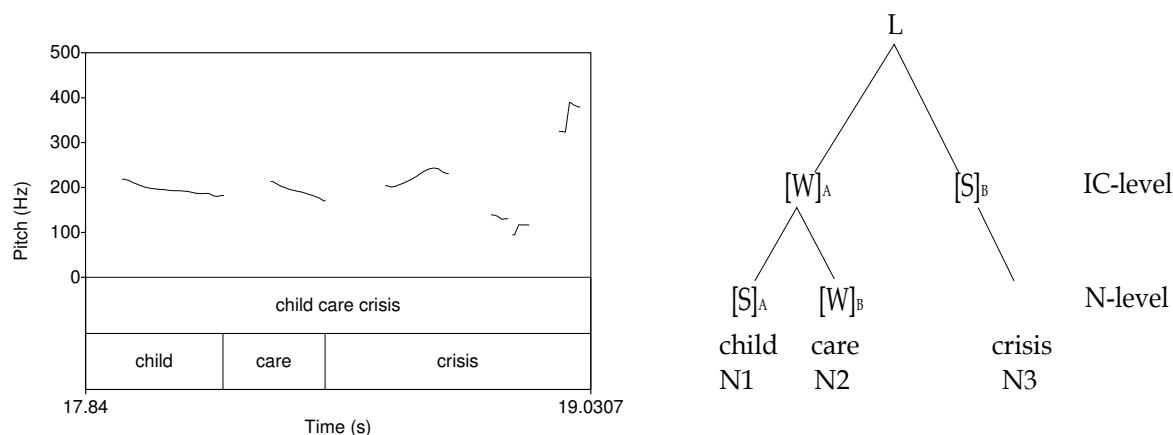
Pattern	P1P2	P2P3	P1P3	Most prominent	# of items	Level
1	IRRELEVANT	NEGATIVE	NEGATIVE	N3	64	IC
2	NEGATIVE	POSITIVE	IRRELEVANT	N2	43	N

constituent N2 is irrelevant in these compounds, and thus among the 64 compounds with a negative P2P3 and P1P3 difference there may be some that have a positive P1P2 difference and others with a negative P1P2 difference. Pattern 2 is exhibited by compounds in which N2 is most prominent, due to a negative P1P2 difference, indicating the prominence of N2 vis-à-vis N1, and a positive P2P3 difference, indicating the prominence of N2 vis-à-vis N3. The pitch difference P1P3 is irrelevant in these compounds as the relation between constituent N2 and constituent N3 is that of left prominence. As noted in table 4.3, this pattern indicates the N-level violations among the left-branching compounds. We can see from the figures that a non-negligible proportion of 35.54% of our left-branching compounds (107 of 301) clearly violate the LCPR, 14.29% at the N-level, 21.26% at the IC-level. For a better illustration, Figure 4.5 shows a pitch track of a compound violating the LCPR at the N-level and Figure 4.6 that of a compound violating the LCPR at the IC-level. The respective prominence pattern of the two compounds is additionally illustrated with a metrical tree.

Figure 4.5: Pitch track and metrical tree of the left-branching compound *science fiction shocker*



Let us now turn to the group of right-branching compounds in more detail in order to find out more about potential violations among that group. As violations of the LCPR I consider all right-branching compounds with highest prominence assigned to constituent N1 or to constituent N3. Whereas highest prominence on constituent N3

Figure 4.6: Pitch track and metrical tree of the left-branching compound *child care crisis*

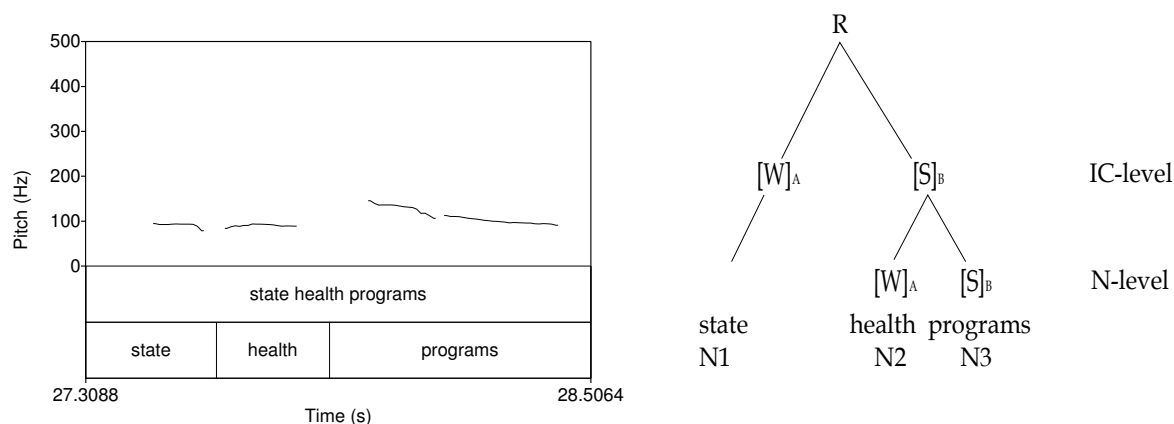
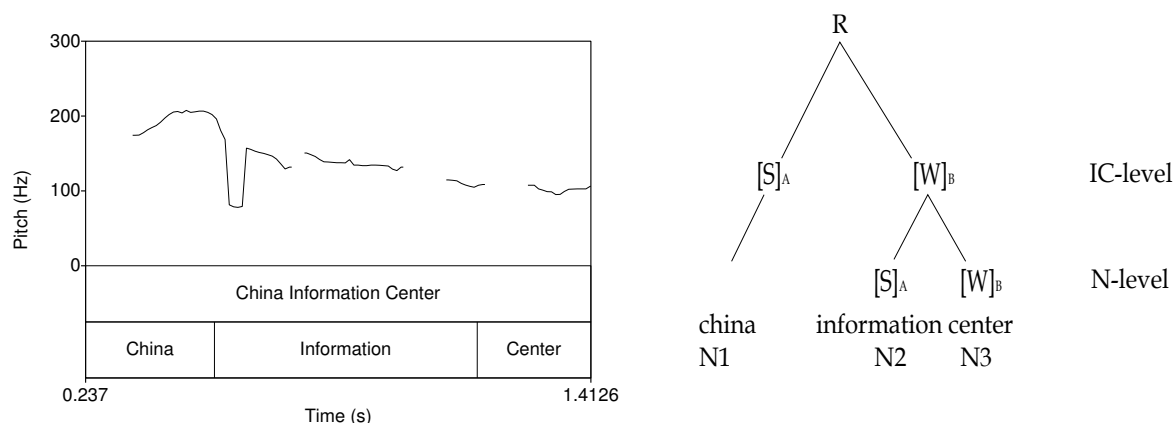
would indicate a violation of the LCPR at the N-level, a right-branching compound with highest prominence on constituent N1 would be a violation of the LCPR at the IC-level. Again, we assume that a negative pitch difference indicates right prominence, whereas a positive pitch difference of more than 2 ST indicates left prominence. Table 4.4 lists the violating patterns for right-branching compounds.

Table 4.4: Violations of the LCPR: right-branching compounds

Pattern	P1P2	P2P3	P1P3	Most prominent	# of items	Level
1	IRRELEVANT	NEGATIVE	NEGATIVE	N3	12	N
3	POSITIVE	positive	POSITIVE	N1	23	IC

Compounds with negative P2P3 and P1P3 pitch differences exemplify exceptions to the rule at the N-level, since for them N3 is most prominent. The pitch difference between constituent N1 and constituent N2 is irrelevant in these compounds as the negative P2P3 and P1P3 differences already indicate that N3 is the most prominent constituent. Hence, compounds showing pattern 1 may have either positive or negative P1P2 differences. The pitch track of a right-branching compound with pattern 1 and the corresponding metrical tree are given in Figure 4.7.

Right-branching compounds with positive pitch differences throughout violate the LCPR at the IC-level. Hence, highest prominence is assigned to constituent N1. However, in Table 4.4 only the P1P2 and P1P3 differences are larger than 2 semitones, which is indicated by the capital letters in the label POSITIVE. The 'positive' label of the P2P3 pitch difference indicates that compounds showing pattern 3 have a positive P2P3 difference, yet which is not necessarily larger than 2 ST. The reason for this is that as soon as the differences P1P2 and P2P3 are both positive, the P1P3 difference

Figure 4.7: Pitch track and metrical tree of the right-branching compound *state health programs*Figure 4.8: Pitch track and metrical tree of the right-branching compound *China information center*

is also automatically positive as it is the sum of the former two. Thus, since the P1P2 difference in these compounds is larger than 2ST, the difference between constituent N1 and constituent N3 is also automatically larger than 2ST. It follows from this that constituent N1 is more prominent than constituent N2 and constituent N3. An example is given in Figure 4.8. We find a total of 30.9% (i.e. 35 out of 113) right-branching compounds that violate the LCPR, with 10.6% violations at the N-level and 20.35% at the IC-level.

We may ask the question as to whether there is a difference between left- and right-branching compounds with regard to their conformity to the LCPR. A Chi-squared test revealed that left- and right-branching compounds do not significantly differ in the proportion of violations found for both groups (Chi-square = 1.13, $p = 0.56$).

Finally, the closer exploration of the data seems to support the previously stated

assumption that the general tendency of left-branching compounds to have a higher pitch on constituent N3 than right-branching compounds may indeed be due to a considerable number of left-branching compounds with highest prominence on constituent N3. The exploration of the data showed that 21.6% of the left-branching compounds exhibited highest prominence on constituent N3. In contrast, we determined only 11.5% right-branching compounds in which N3 was the most prominent constituent.

In sum, the exploration of potential violations of the LCPR has revealed that there seems to be a substantial number of compounds that violate the LCPR, and these violations occur at both IC- and N-levels, and in both left- and right-branching compounds.

4.3 Discussion

The analysis of the prominence patterns of triconstituent compounds has provided empirical support for Liberman and Prince's (1977) Lexical Category Prominence Rule. Testing the predictions made by the LCPR, I found that left-branching compounds tend to have highest prominence on constituent N1 whereas right-branching compounds tend to have highest prominence on constituent N2. However, as shown in particular in section 4.2.2, I also found a considerable number of compounds that do not behave according to the LCPR. About one third of both left- and right-branching compounds belong to this ill-behaved group. Yet, the result that some left-branching compounds exhibited highest prominence on constituent N2 and constituent N3, respectively, is in line with previous findings, for example, by Berg (2009), Kvam (1990) and Giegerich (2009). The exceptions found for right-branching compounds also match findings by Berg (2009) and Giegerich (2009) and illustrate that even in the presence of a complex head, prominence assignment in triconstituent compounds is quite variable. Yet, the question remains of what causes the violations of the LCPR at the N-level and IC-level, respectively. Let us first have a closer look at some of the N-level violations and then turn to the IC-level violations.

4.3.1 Violations of the LCPR at the N-level

As illustrated in the previous subsection (see Figure 4.5), the pitch analysis of the left-branching compound *science fiction shocker* revealed that the compound exhibits highest prominence on constituent N2, which is a violation of the LCPR at the N-level.

In section 2.2 I have suggested that such N-level violations of the LCPR may be triggered by right prominent NN compounds, which occur embedded in triconstituent compounds and whose prominence pattern is preserved under embedding. The possibility that right prominent NNs occur embedded in larger compounds is ignored in the formulation of the LCPR as it is built on the assumption that compounds are generally left prominent. Yet, according to various dictionaries (e.g. *Longman Dictionary of American English*, *Oxford's Student Dictionary of American English*, *Longman Advance American Dictionary*) the embedded NN compound *science fiction* is in fact right prominent. Thus, as argued in section 2.2, we find empirical evidence that in *science fiction shocker* the prominence pattern of the right prominent NN compound *science fiction* is preserved under embedding and causes highest prominence on constituent N2 of the compound, contra the LCPR. Additional examples of this kind from the corpus are *Thanksgiving day*, *grass roots advocates*, *capital gains tax* and *home improvement loans*. According to the acoustic analysis, in all of these compounds highest prominence is clearly assigned to constituent N2, and their complex constituent is also attested as right prominent in various dictionaries. Furthermore, among this subset of violations we find compounds such as *governor-sergeant appointee*, *Mattapan-Roxbury area* and *felony-sodomy charges*, whose complex constituents belong to the class of copulative compounds. This class is uncontroversially considered to be right prominent (e.g. Fudge 1984; Olsen 2000; Plag 2003), and, like in the other cases, their prominence pattern is preserved under embedding.

The same kind of prominence preservation is also observable for right-branching compounds, although the number of examples is quite small. In fact, there are only two compounds in this subgroup for which I am quite confident that the embedded NN is in fact right prominent, i.e. the compounds *Operation desert storm* and *Operation desert shield*. In both of these compounds the complex constituent forms a proper noun, which is another class of compounds that was found to show a significant statistical tendency towards right prominence in the study by Plag et al. (2008). With reference to the other 10 right-branching compounds violating the LCPR at the N-level, I rather remain agnostic with respect to the prominence pattern of the embedded NNs. The reason for this is that I did not find the embedded NNs attested with neither left- nor right prominence in any of the dictionaries and corpora I consulted as control sources.

In sum, it appears to be the case that a large portion of the N-level violations in the corpus can be explained by the fact that, contra the LCPR, a considerable number of the embedded compounds are right prominent, and that this right prominence is

preserved under embedding, just like left prominence is preserved under embedding according to the LCPR.

4.3.2 Violations of the LCPR at the IC-level

In addition to the violations at the N-level, the corpus also contains some left- and right-branching compounds violating the rule at the IC-level. In these compounds highest prominence is assigned to the constituent outside the complex constituent, i.e. constituent N3 in left-branching compounds and constituent N1 in right-branching compounds. However, an explanation for the violations at the IC-level seems less clear than the one I offered for the N-level violations. As argued in section 2.2, one possibility would be that left-branching IC-level violations might be caused by the same factors that trigger prominence variation in biconstituent compounds (N-level). According to studies by Plag et al. (2008), Plag and Kunter (2010) and Kunter (2011), these factors might be semantics and analogy. In fact, a closer look at the group of left-branching compounds violating the LCPR at the IC-level suggests that their prominence pattern may be indeed explained by means of semantics. In particular, I found a number of left-branching compounds which uncontroversially showed one of these semantic relations that turned out to trigger right prominence in NN compounds. For instance, the compounds *Bay state voters*, *Boston area communities*, *Beacon Hill democrats*, *nursing home patient* and *weekend séries* revealed a locative and temporal relationship, respectively, at the IC-level. Furthermore, I detected compounds with the semantic relation *IC1 has IC2* (e.g. *waste company officials*, *state lottery officials*, *oil company executive*, *Beacon Hill insiders*), which is also one of the semantic relations that was found to trigger right prominence in NN compounds in the study by Plag et al. (2008).¹⁶ However, as I have not systematically coded the corpus data neither for semantic relations nor semantic categories, this assumption requires further empirical support. This also applies to the factor analogy with respect to which I cannot make any reliable statements based on the present data. Yet, as mentioned in section 2.2, I specifically tested the potential effect of semantics and analogy on the prominence behaviour of triconstituent compounds in a production experiment using carefully controlled experimental data (see chapters 5 and 7 for a description of the experiment and its results).

Thus, let us turn to right-branching compounds violating the LCPR at the IC-level, i.e. compounds in which constituent N1 was marked as the most prominent con-

¹⁶It should be noted here that for some compounds there is more than one interpretation, as for instance, the compound *nursing home patient* may be interpreted as *IC1 has IC2* as well as *IC2 is located at/in IC1*.

stituent. These compounds are left prominent at the IC-level despite the presence of a right-hand branching node. Hence, such compounds provide evidence that even in the presence of a complex head, the LCPR fails to predict the correct prominence pattern. However, if it is not branching direction that governs prominence assignment in these compounds, what else might be responsible for these violations? One plausible answer to this question may be that the same factors thought to trigger right prominence at the IC-level of triconstituent compounds also operate in the opposite direction, i.e. causing left prominence. For instance, there might be semantic relations or categories that trigger left prominence at the IC-level, which would override the potential effect of branching and lead to highest prominence of constituent N1. For instance, Plag et al. (2007) found a particularly strong statistically significant tendency towards left prominence for the semantic relation *N2 for N1*. Support for the idea that this effect may be responsible for some of the LCPR violations comes from Giegerich (2009). He provides some examples of right-branching compounds with highest prominence on constituent IC1, namely *tómato green-house*, *gráin store-room*, *stéel ware-house*, *ówl nest-box*. All of these compounds can be paraphrased as *N2 for N1*, the semantic relation which seems to trigger left prominence. With regard to our corpus, this explanation may hold for the compounds *cómmunity meeting hall* ('a meeting hall for the community') and *crédit scoring system* ('a scoring system for credit').

Bringing in semantics may also lead to a complete abandonment of branching direction as a factor in prominence assignment. Thus, one could argue that if the semantic relation at the IC-level triggers right prominence, main prominence on IC2 is not due to branching, but due to the semantics at the IC-level. This is in the spirit of Selkirk (1984), who claims that right-branching compounds are right prominent because the relation between the complex and single constituent is always that of a modifier-head relation (which is taken to trigger right prominence) but never that of an argument-head relation (which would trigger left prominence). From recent research (e.g. Plag 2006; Plag et al. 2007; Plag et al. 2008; Kunter 2011) we know, however, that it is only certain semantic relations and categories that favour right prominence, and not all modifier-head relationships.

Now, among the right-branching compounds I find that a majority of compounds with highest prominence on IC2 exhibit those semantic relations which are claimed to trigger right prominence in biconstituent NN compounds. Hence, in these compounds, both branching and semantics would favor right prominence at the IC-level. A list of a few examples is given in Table 4.5.

Apart from semantics there might be yet another possible explanation for some of

Table 4.5: Semantic relations at the IC-level of some right-branching compounds with highest prominence on IC2

Compound	Semantic relation
<i>brick townhouse</i>	IC2 is made of IC1
<i>Iowa cornfield</i>	IC2 is located at/in IC1
<i>corner drug store</i>	IC2 is located at/in IC1
<i>School drug use</i>	IC2 is located at/in IC1
<i>Rockingham horse track</i>	IC2 is located at/in IC1
<i>Yale law school</i>	IC2 has IC1
<i>State taxpayers</i>	IC1 has IC2
<i>Roxbury housing project</i>	IC1 has IC2
<i>Hynes convention center</i>	IC2 is named after IC1

the violations, i.e. the influence of “pragmatic interpretive strategies” (Ladd 1984) on the prominence patterns of compounds. Ladd (1984) argues that the head of a compound may be deaccented if it is semantically not very specific. This may be the case for the compound *China information center* (see Figure 4.8 in the previous section), which exhibits highest prominence on constituent N1. With reference to this compound, we may argue that the crucial information is conveyed by the constituent on the left, i.e. *China* rather than on the complex head *information center*. The same explanation might be valid for a number of other compounds found among the violations (e.g. *Washington law professor*, *Superman comic book*). Although Ladd does not explicitly refer to triconstituent compounds in his discussion, there is no reason why this assumption should not also hold for more complex compounds.

Finally, among the ill-behaved group there are also a few compounds whose heads are strongly lexicalized with a high token frequency (e.g. *Fox network*, *traffic headaches*, *Tiffany network*, *Massachusetts congressman*, *Logan airport*). Given that items with a higher token frequency are more expectable, hence less informative, than items with a lower token frequency, these constructs could also be explained in terms of Ladd’s deaccentuation account.

In sum, the IC-level violations provide evidence that the factor branching direction alone does not govern prominence assignment in triconstituent compounds. This implies that other factors must play a role as well. These factors may be for instance semantics or information structure, for which there is some evidence in the data.

4.4 Summary and conclusion

This chapter described a corpus study which tested the predictions of Liberman and Prince's (1977) Lexical Category Prominence Rule for left- and right-branching nominal compounds in English. The study provides empirical evidence that the majority of left- and right-branching compounds behaves as predicted by the LCPR. Yet, a more detailed exploration of the data also discovered a significant proportion of left- and right-branching compounds not behaving according to Liberman and Prince's rule. Among these exceptions, we distinguished between compounds violating the LCPR at the N-level and those violating the rule at the IC-level of the compounds. The compounds classified as N-level violations are left- and right-branching compounds with highest prominence on the right member of the complex constituent instead of the predicted left one (e.g. [*science fiction*] *shocker*, [*grass roots*] *advocates*, [*capital gains*] *tax*, *state* [*health program*] etc.). In contrast, the IC-level violations are left- and right-branching compounds with highest prominence on the constituent outside the complex constituent, i.e. on constituent N3 and on constituent N1, respectively (e.g. [*child care*] *crisis*, *China* [*information center*]). The IC-level violations suggest that a complex head constituent, i.e. a right-hand branching node, does not necessarily trigger right prominence nor that the absence of a complex head constituent automatically causes left prominence. This result is in line with the result of previous studies by Kvam (1990), Berg (2009) and Giegerich (2009).

In general, the N- and IC-level violations determined in the corpus study strongly suggest that it is not necessarily the branching direction of the compounds that governs prominence assignment in triconstituent compounds but that other factors must play a role. Among the N-level violations there are some compounds which provide first empirical support for the assumption that right prominent NN compounds occur embedded in larger compounds due to which highest prominence is assigned to the right member of the complex constituent instead of the left one. The IC-level violations provide some first hints that semantics and informativeness may be influential factors of prominence assignment in triconstituent compounds. Yet, as mentioned before, the corpus data was neither systematically controlled for the prominence pattern of the embedded NN nor the semantic relations between the individual compound constituents. Hence, these assumptions should be tested again with more carefully controlled data as it is done in the two production experiments presented in the following three chapters. The experiments take a closer look at right prominent NN compounds and their potential influence on the prominence behaviour of triconstituent

compounds. In particular, I address the question whether such compounds affect prominence assignment in triconstituent compounds, either directly through embedding or indirectly due to the same factors responsible for rightward prominence in NN compounds, i.e. semantics and analogy.

5 Experimental data

5.1 Introduction

As argued in section 2.2, the existence of right prominent NN compounds is ignored in the formulation of the LCPR. However, as shown in various experimental and corpus studies (e.g. Plag 2006; Plag et al. 2007; Lappe and Plag 2007; Plag et al. 2008; Kunter 2011) right prominent NN compounds do exist in English. Given that the LCPR is based on the assumption that compounds are generally left prominent, the existence of such right prominent NNs must also have serious consequences for prominence assignment in triconstituent compounds. For instance, apart from embedded left prominent NNs we may also find right prominent NNs to occur embedded in triconstituent NNN compounds. This assumption and the assumption that prominence is generally preserved under embedding (e.g. Liberman and Prince 1977) may cause triconstituent NNN compounds to have highest prominence on the right member of the complex constituent instead of the left one (see also Giegerich (2009) on that matter). Thus, right prominent NN compounds may directly affect prominence assignment in triconstituent compounds and some first empirical support for this assumption has already been provided by the corpus study presented in the previous chapter.

In addition to that, it was argued in section 2.2 that the same factors said to trigger right prominence in NN compounds may also be responsible for rightward prominence at the IC-level of left-branching NNN compounds (i.e. N3). Thus far, such compounds have been simply referred to as being phrases by most authors, with their prominence pattern said to be governed by the ‘Nuclear Stress Rule’. However, as argued earlier in this thesis, this approach is highly questionable due to the lack of convincing criteria that would support such a phrasal analysis of those left-branching compounds. However, given that triconstituent compounds are also binary in nature and prominence in triconstituent compounds is said to be governed by the same rule that governs prominence in NN compounds, it may be possible that the same factors operating in NN compounds also operate in NNN compounds. Hence, the existence of right prominent NNs may be said to be indirectly responsible for the aberrant

prominence behaviour to the LCPR.

These potential effects, which the existence of right prominent NN compounds may have on prominence assignment in larger compounds, are tested with two production experiments which are described in this and the following two chapters. Production experiment 1 focuses on the prominence pattern of the embedded NN compounds and their influence on prominence assignment to left- and right-branching NNN compounds. It is tested whether left- and right-branching NNN compounds with embedded left prominent NNs have highest prominence on constituent N1 and constituent N2, respectively, i.e. according to the LCPR. In addition to that, it is investigated whether left- and right-branching NNN compounds with embedded right prominent NNs have highest prominence on the right member of the complex constituent, i.e. constituent N2 in left-branching compounds and constituent N3 in right-branching compounds, contra the LCPR. These two predictions have been subsumed under what I referred to as the Embedded Prominence Hypothesis (EPH) in section 2.2. The EPH is again given in 1.

1. **Embedded Prominence Hypothesis (EPH):**

If highest prominence falls on the complex constituent of a triconstituent compound:

- a) left- and right-branching compounds with an embedded left prominent NN compound have highest prominence on the left member of the complex constituent, i.e. N1 and N2 respectively.
- b) left- and right-branching compounds with an embedded right prominent NN compound have highest prominence on the right member of the complex constituent, i.e. N2 and N3 respectively.

Production experiment 2 focuses on the question as to whether the same factors assumed to trigger right prominence in NN compounds, namely semantics and analogy, also trigger right prominence at the IC-level of left-branching NNN compounds. This hypothesis was referred to as the IC-Prominence Hypothesis (IPH) in section 2.2 and is given again in 2.

2. **IC-Prominence Hypothesis (IPH):** The same factors thought to trigger right prominence in biconstituent NN compounds, i.e. semantic and analogical factors, also trigger right prominence at the IC-level of left-branching NNN compounds.

In the remaining of this chapter, I provide some details on the speakers participating in the two production experiments as well as on the general experimental design. Furthermore, I provide information on the methodology common to both experiments,

which includes a description of the recording procedure and the subsequent annotation process of the data with the speech analysis PRAAT.

5.2 Common methodology: Experiments 1 and 2

5.2.1 Subjects

The two experiments were conducted with a total of 36 speakers of North American English at the University of Toronto. The only prerequisite to participate in the experiment was to be a native speaker of North American English. No further restrictions were made with regard to whether the speakers were American or Canadian nor the regional variety spoken by the participants. Potential influences introduced by regional differences were thought to be accounted for by devising a mixed-effects model with the speakers modelled as random effect (see section 3.4 again). Furthermore, bilingual speakers did not qualify as subjects for the two experiments. The reason to exclude bilingual speakers was to minimize additional sources of variation as possibly introduced by a second mother tongue.¹

The 36 speakers participating in the two experiments were undergraduate and graduate students, respectively, aged between 17 to 40.² The speakers were told that the experiment is on the pronunciation of North American English. Data of five speakers, i.e. four speakers participating in experiment 1 and one speaker participating in experiment 2, had to be entirely discarded before the data analysis due to the speakers' struggle to read the test sentences fluently or in a natural manner. In addition, one speaker became aware of the focus of the experiment during the recording procedure, which affected his pronunciation of the relevant compounds. In particular, during the recording procedure, the speaker started to read the sentences with the embedded NNN compounds twice, each time with a different prominence pattern assigned to the respective compound. Because of that, the data of this particular speaker was excluded from the final data set. Of the remaining 13 speakers participating in experiment 1, six speakers were male and seven speakers were female. Eight speakers grew up in the province of Ontario whereas four speakers had lived most of their lives in other provinces of Canada before moving to Ontario. One speaker was originally from Massachusetts, USA, who had lived in Toronto for two years at the time of the

¹In this thesis, I refer to bilingual speakers as speakers who simultaneously acquired two languages as a child. This was made explicit to the participants when they filled out the questionnaire in the beginning of the recording session.

²The majority of the participating speakers was between 18 and 25 years old.

recording. In experiment 2, ten speakers were female and seven speakers were male; 13 speakers grew up in the province of Ontario whereas two speakers were originally from other Canadian provinces. One speaker was from Detroit, Michigan (USA) and one speaker did not provide any information regarding his birth place, yet he provided the information of being Canadian. According to the subjects none of them had any hearing or speaking disorders.

5.2.2 Experimental design

The two experiments were designed as reading experiments in which the participants were asked to read out aloud lists of sentences while being recorded. The lists presented to the participants contained on the one hand sentences with embedded NNN compounds, i.e. the constructed stimuli, and on the other hand about the same number of fillers. The filler sentences differed from the sentences containing the compounds in that they contained no triconstituent Noun+Noun+Noun compound. The fillers were added to the test sets in order to distract the speakers from the actual aim of the experiment. The data set for experiment 1 consisted of 80 sentences, i.e. 40 sentences containing a compound and 40 fillers. The data set of experiment 2 consisted of a total of 88 sentences, i.e. 44 sentences with embedded NNN compounds and 44 filler sentences. As for the the left- and right-branching NNN compounds itself, the construction of the sentences in which the compounds were embedded underlied certain criteria, which are described in more detail in the following paragraph.

The 40 compounds constructed for experiment 1 and the 44 compounds constructed for experiment 2 were embedded in simple declarative sentences which served as carrier sentences for the compounds. I embedded the compounds in carrier sentences rather than presenting them in isolation because I wanted to test the two hypotheses for compounds embedded in a relatively natural speech context. In order to resemble such a natural speech context, I adapted the context of each carrier sentence to the semantics of the embedded NNN compounds. Thus, instead of using the same standard carrier sentence for all compounds, I constructed individual carrier sentences for each compound.

Furthermore, although I mentioned in section 3.3 that the sentence position of a compound does not seem to have a crucial effect on its prominence pattern (e.g. Plag 2006), I decided to control for this factor while constructing the sentences. Thus, the sentence position of each compound was held constant in all carrier sentences, with each compound being placed in object position of its carrier sentence. In the filler sen-

tences of the two subsets, the object slot was filled by a long noun phrase.³ In addition, I was also consistent with reference to the number of sentence constituents in each carrier sentence in order to minimize additional sources of variation.⁴ Besides, I controlled for contrastive stress environments and information status of the compounds. As it was already mentioned in section 3.3, contrastive stress environments such as “He read about a coffee **table** designer not a coffee **mug** designer” may lead to a change of the canonical pattern of a compound (e.g. Bauer 1998). Hence, such environments were avoided when constructing the sentences of the two experiments. Furthermore, it was decided to present all compounds as discourse-new information as it has been mentioned in the literature that discourse-old information may lead to a deaccentuation of the head constituent of a given compound (e.g. Hirschberg 2002). Finally, all embedded NNN compounds were followed by a two-word adverbial, which at the same time marked the end of each carrier sentence. The reason to add this adverbial to the end of each sentence was to avoid boundary tones to affect the automatic pitch measurements. Boundary tones are said to occur at the edge of intonational phrases and are part of the intonational contour of an utterance (cf. Ladd 1996:80). Hence, these boundary tones are not part of the prosodic structure of a compound and thus they should be controlled for in an analysis that only relies on pitch measurements in order to account for prominence differences in compounds.

For a better illustration of the created test and filler sentences a few examples are given in (1) and (2) (see part A of the Appendix for the full test set). The sentences in (1) are the test sentences whereas those in (2) are examples of filler sentences.

- (1) *She started **hay fever treatment** last year.*
*She founded a **student string orchestra** last month.*
- (2) *She attended a **Spanish and French class** last semester.*
*She missed her **favourite TV show** last night.*

The test sets of experiment 1 and experiment 2, respectively, were presented to the participants in one of three different orders. Altering the order of the sentences was thought to avoid sequencing effects as possibly caused by fatigue of the readers. The first set of each experiment was a pseudo-randomization with the other two sets con-

³It should be noted that some of the noun phrases in the filler sentences contained NN compounds as their head constituent.

⁴The number of constituents in the filler sentences sometimes differed from that of the sentences with the embedded compounds. Crucially, the sentences with the embedded compounds all shared the same number of sentence constituents.

structured by dividing the total number of sentences into smaller blocks of 10 sentences, which were then systematically varied. The test sets were presented to the participants on different sheets of papers.

5.2.3 Recording procedure

The recordings took place in a sound proof booth at the University of Toronto. All recordings were carried out with the portable solid state audio recorder Marantz PMD660 which records in digital audio formats onto Compact Flash memory cards. The readings were recorded with an external microphone placed in front of the speakers before the beginning of each recording session. The readings were recorded as wave files using a 44.1 kHz sample rate. Before each recording session, the subjects were asked to fill out a short questionnaire and to read through the instructions provided to them on a piece of paper. In addition, the subjects were instructed to read five sentences not belonging to the original test set. These sentences differed from the sentences of the experiment in that they contained no compounds (see part A of the Appendix for the full test sets). These training sentences were necessary for the researcher to adjust the recorder to the subjects' individual voices, i.e. the loudness level, and to familiarize the subjects with the test situation as well as with the process of reading out loud. The regular test sets were then presented to the subjects on different sheets of paper placed on the table in front of them. Before each recording, the subjects were explicitly instructed to read the test sentences as naturally as possible and to repeat those sentences, which they felt they had not read fluently or in a natural manner. Each recording took about 15 to 25 minutes depending on the subjects' reading performance. The researcher remained in the sound proof booth during the recording procedure, and asked subjects to repeat sentences, in which they stuttered, hesitated or mispronounced words.

5.2.4 Acoustic measurements

The prominence pattern of the experimental data was determined by means of pitch analysis. Following the method employed in the corpus study, in PRAAT textgrids, each compound, each compound constituent as well as each sonorant part of the rime of the highest prominent syllable in each constituent was manually segmented. Particular attention was paid to occurrences of prominence shifts, (e.g. *campáign + mánager* becomes *cámpaign manager*) which were segmented accordingly. The segmentation of the compounds was done on the basis of a combined waveform and spectrogram

view. Furthermore, all recordings were manually checked again as to whether the sentences were read fluently and in a natural manner, i.e. without list intonation. Furthermore, it was checked whether in those sentences in which the relevant compounds were embedded, the speakers actually pronounced all constituents. Sentences that were not fluently read or in which constituents were left out were excluded from the analysis. This resulted in the exclusion of 4 tokens from the data set of experiment 1 and 27 tokens from that of experiment 2, which reduced the total number of tokens to 516 and 716 tokens, respectively.

The same automatic pitch tracking algorithm that has been used in the corpus study was also applied to measure the mean pitches of the experimental data. Hence, the algorithm automatically adjusted to octave jumps, creaky voice phonation and the speakers' individual pitch range.⁵ The pitch algorithm returned 1522 pitch values for the data of experiment 1 and 2129 pitch values for the data of experiment 2. For 16 observations of experiment 1 and 19 observations of experiment 2, however, the algorithm failed to return a pitch value, despite its various automatic adjustments. A manual checking of these recordings showed that the algorithm's failing was either due to the shortness of the relevant measurement interval or due to strong creaky voice phonation.⁶ Moreover, measurement errors caused by portions of non-modal phonation and background noises, respectively, were manually corrected. This was done by dividing the total number of glottal pulses of the respective measurement interval by its duration.

Finally, I excluded another 43 pitch values from the data set of experiment 1 and 35 pitch values from the data set of experiment 2 because of the presence of strong creaky voice phonation during those speech segments. As it was already discussed in detail in section 4.1.3, creaky voice is associated with unnaturally low F0 values, which are caused by irregular vibrations of the vocal cords. Given that we rely on pitch as the only cue to prominence, these values had to be excluded in order to avoid them to affect the final pitch analysis. This reduced the total number of pitch observations to 1479 in experiment 1 (i.e. 516 tokens) and 2094 in experiment 2 (i.e. 711 tokens).⁷ All

⁵See chapter 4.1.3 again for more details on the algorithm's specific pitch settings.

⁶The observations for which the algorithm failed to measure a pitch value belonged primarily to items whose nucleus in the most prominent syllable was the short front vowel [i] as for instance in *clinic*, *Christmas*, *city* or *ticket*. For these segments the interval length was less than two cycles long, which is the minimum length required by a pitch algorithm to calculate the fundamental frequency of the speech signal (cf. Ladefoged 2003:77).

⁷Whereas the exclusion of the 43 pitch values from the data set of experiment 1 did not affect the total number of tokens, the exclusion of the 35 pitch values from the data set of experiment 2 reduced the total number of tokens from 716 to 711. Furthermore, due to the exclusion of these pitch values, not all compounds in the two data sets had three pitch values, but for some of them only one or two

remaining pitch values were transformed into semitones relative to the lowest pitch value in the respective data set (see section 4.1.3 again for the respective formula).

values were available.

6 Production Experiment 1: Embedded Prominence Hypothesis

This chapter presents a production experiment that tests the predictions of the Embedded Prominence Hypothesis as developed in section 2.2. Section 6.1 focuses on some methodological aspects, which are specific to the current experiment. The section provides information on the stimuli constructed in order to test the EPH, introduces the predictions for the pitch analysis and provides details on the statistical modelling of the data. Section 6.2 presents the results of the experiment, which are further discussed in section 6.3. Section 6.4 deals with data obtained for one particular speaker, whose pitch values were notably different from that of the other speakers. The chapter ends with a conclusion in section 6.5.

6.1 Methodology

6.1.1 Stimuli

In order to test the Embedded Prominence Hypothesis for compounds with embedded left prominent NNs, 20 left- and right-branching compounds (10 compounds for each group) with embedded left prominent NN compounds were constructed. In addition to that, 20 left- and right-branching compounds with embedded right prominent NN compounds were created to test whether the presence of a right prominent complex constituent directly affects the overall prominence pattern of a triconstituent compound as predicted by the Embedded Prominence Hypothesis. For a better illustration of the four created compound groups, an example of each group is given in Table 6.1.

The prominence pattern of each embedded NN compound was controlled by means of various American English dictionaries (*Oxford Student's College Dictionary of Amer-*

Table 6.1: Types of compounds constructed to test the EPH

branching direction	Prominence pattern of embedded NN	NNN compound
left-branching	left prominent [N̂N]	[N̂N]+N
right-branching	left prominent [N̂N]	N+[N̂N]
left-branching	right prominent [N̂N̂]	[N̂N̂]+N
right-branching	right prominent [N̂N̂]	N+[N̂N̂]

ican English (Hornby 1983), the *Longman Dictionary of American English* (Bullon 2002), the *Longman Advanced Dictionary of American English* (Summers 2000), and the *Oxford Advanced Learner's Dictionary* (Hornby 1995).¹ Furthermore, I used a list of compounds provided by Teschner and Whiteley (2004).² Only NN compounds whose prominence pattern was attested in at least one of the above mentioned sources qualified as potential complex constituents for the triconstituent compounds. Yet, compounds for which the above mentioned sources differed in the assigned prominence pattern did not qualify as complex constituents. The exclusion of these compounds was crucial in order to reduce the risk of introducing NN compounds with a variable prominence pattern to the data set. Due to the same reason, I also avoided selecting compounds that have been generally argued to show a variable prominence behaviour (e.g. *ice cream* vs. *ice créam* (Bloomfield 1933:228)).

Apart from controlling for the prominence pattern of the embedded NN compounds, testing the Embedded Prominence Hypothesis also required to control for various factors at the IC-level of the constructed triconstituent compounds. First, it was crucial that the constructed compounds were clearly left- and right-branching in order to exclude the possibility that potential prominence variation would be caused by structurally ambiguous compounds. As already mentioned in section 4.1.2, the problem with structurally ambiguous compounds such as *kitchen towel rack* or *silver knife handle* is that highest prominence is claimed to be either assigned to constituent N1 or constituent N2, depending on the compounds' interpretation (e.g. Warren 1978; Visch 1999). For example, the compound *kitchen towel rack* is structurally ambiguous in that it may be interpreted as either 'a rack for a kitchen towel' (left-branching) or 'a towel rack located in the kitchen' (right-branching). In case of a left-branching in-

¹The edition of the *Oxford Advanced Learner's Dictionary* used in this thesis, came with a CD-ROM. The CD-ROM provided information about both the British and the American pronunciation of its entries and thus became an appropriate control source for the NN compounds.

²Teschner and Whiteley (2004) is a textbook on English pronunciation for college students of English. The textbook comes with a CD-ROM, which contains, among other things, a list of compounds and phrases extracted from the *Oxford Spanish-English Dictionary* (Carvajal and Horwood 1996).

terpretation, the LCPR assigns highest prominence to constituent N1. In contrast, if the compound is interpreted as right-branching, highest prominence is assigned to constituent N2 by the rule. Given this variable prominence behaviour of structurally ambiguous compounds and given the fact that it is difficult to control which interpretation the speakers may choose, I tried to avoid the construction of such ambiguous compounds.

Second, as the present experiment particularly focused on the effect of the embedded NN compound, I also tried to control for the prominence relation at the IC-level of the constructed left-branching compounds. For left-branching compounds with embedded left prominent NNs, the EPH predicts that highest prominence is assigned to constituent N1, whereas for left-branching compounds with embedded right prominent NNs highest prominence is expected to be assigned to constituent N2. Hence, in order to test the EPH, i.e. the effect of the embedded NN, it was necessary to construct left-branching compounds which would be left prominent at the IC-level, i.e. in which highest prominence would not fall on constituent N3. But how to control for the prominence relation at the IC-level of these compounds without knowing the factors that may cause right prominence at that level?

One chance to minimize the risk of right prominence at the IC-level of these compounds was to control for at least those factors that had been found to trigger right prominence in biconstituent NN compounds - although at the time of the experiment it was still unclear whether these factors were in fact relevant. Hence, the left-branching compounds were constructed in such a way that they would not exhibit one of those semantic relations found to trigger right prominence in NN compounds, as for instance, *IC2 is located at IC1* (e.g. [*Boston area*] communities), *IC2 is during IC1* (e.g. [*weekend*] series) or *IC2 is made of IC1* (e.g. [*chocolate cream*] cake). Furthermore, I also avoided to select items that would belong to one of those semantic categories that turned out to trigger right prominence in NN compounds in studies by Plag et al. (2007), Plag et al. (2008) and Kunter (2011). Examples of such categories are, *IC1 is a proper noun* (e.g. [*Wall Street*] journal), *IC1 is a temporal modifier* (e.g. [*morning news*] interview) or *IC2 is a geographical term* (e.g. [*Sunflower*] valley). In addition to that, I avoided selecting as head constituents lexical items such as *avenue* and *pie*, which have been generally claimed to trigger right prominence in NN compounds (e.g. Ladd 1984; Liberman and Sproat 1992; Bell 2008). By applying the above mentioned criteria, the Embedded Prominence Hypothesis, i.e. the role of the embedded NN compound, could be tested without any of these factors potentially intervening at the IC-level of the left-branching compounds.

But what about right-branching compounds? Regarding the group of right-branching compounds, it was not necessary to exclude the above mentioned semantic relations at the IC-level. Quite to the contrary, right branching compounds had to be right prominent at the IC-level in order for the predicted effect of the embedded NN to surface. Hence, I have not controlled for the semantic factor in right-branching compounds as done for left-branching compounds.

The final data set comprised a total of 40 compounds, i.e. 10 compounds for each of the four compound types. The compounds are abbreviated ‘L/N1’, ‘R/N2’, ‘L/N2’ and ‘R/N3’ for the rest of this thesis. The labels ‘L’ (left) and ‘R’ (right) mark the two branching directions at the IC-level of the compounds. The labels ‘N1’, ‘N2’ and ‘N3’, respectively, refer to the three constituents of the compounds: ‘N1’ refers to the leftmost constituent, ‘N2’ to the second constituent and ‘N3’ to the rightmost constituent. In the abbreviations, ‘N1’, ‘N2’ and ‘N3’ label the constituent that the Embedded Prominence Hypothesis predicts to be the most prominent one in a given compound. For example, a left-branching compound with an embedded left prominent NN compound is labelled L/N1 in the present data set, as the Embedded Prominence Hypothesis predicts highest prominence on constituent N1 in such compounds. R/N2 compounds represent the group of right-branching compounds with an embedded left prominent NN compound; in these compounds the most prominent constituent should be constituent N2 according to the Embedded Prominence Hypothesis, hence the label R/N2. Furthermore, the two compound groups with embedded right prominent NN compounds were labelled L/N2 and R/N3, respectively. According to the Embedded Prominence Hypothesis, L/N2 and R/N3 are expected to exhibit highest prominence on the right member of the complex constituent, hence the labels ‘N2’ and ‘N3’. For a better illustration, Table 6.2 provides an example of each compound type together with its respective label (for a full list of the constructed compounds see Appendix part A). The highest prominent constituent is marked by an acute accent on the nucleus of the relevant syllable.

Table 6.2: Example of each compound type constructed to test the EPH

branching direction	embedded NN	NNN compound	label
left-branching	<i>háy fever</i>	<i>[háy fever] treatment</i>	L/N1
right-branching	<i>shéet music</i>	<i>piano [shéet music]</i>	R/N2
left-branching	<i>science fiction</i>	<i>[science fiction] shocker</i>	L/N2
right-branching	<i>Christmas dinner</i>	<i>family [Christmas dinner]</i>	R/N3

6.1.2 Embedded Prominence Hypothesis: predictions

In order to find out whether the Embedded Prominence Hypothesis holds for left- and right-branching compounds with embedded left prominent NN compounds, the average pitch values obtained for the three constituents of L/N1 and R/N2 compounds had to be compared to each other. According to the Embedded Prominence Hypothesis, in left-branching compounds with embedded left prominent NNs (L/N1) highest prominence falls on constituent N1 whereas in right-branching compounds with embedded left prominent NNs (R/N2) highest prominence falls on constituent N2. Hence, under the assumption that the EPH is correct, we would expect to find the following two pitch patterns for the two created compound groups: For L/N1 compounds, we should find a relatively high pitch assigned to constituent N1, with a clearly lower pitch assigned to constituent N2 and constituent N3. In contrast, for R/N2 compounds, we would expect both constituent N1 and constituent N2 to exhibit a relatively high pitch in order to mark constituent N2 to be more prominent than constituent N1. The lowest pitch is expected to be assigned to constituent N3 as this constituent is assumed to be the least prominent constituent in R/N2 compounds.

For a better illustration of the two predicted pitch patterns, the pitch tracks of a left- and right-branching compound that behave according to the Embedded Prominence Hypothesis are provided in Figure 6.1 and Figure 6.2.³ The horizontal axis displays the duration of the compound, the vertical axis gives the pitch range in Hertz (Hz). The line in the middle of the graph represents the pitch curve for each compound as detected by the Praat pitch tracking algorithm.

The pitch track of the left-branching compound [*hay fever*] *treatment*, which is shown in Figure 6.1, clearly illustrates that highest pitch is assigned to constituent N1 in this compound with a much lower pitch assigned to constituent N2 and constituent N3. In the right-branching compound *piano* [*sheet music*] shown in Figure 6.2, we observe a relatively high pitch on constituent N1 and constituent N2, with the pitch on N2 being even slightly higher than on constituent N1. The lowest pitch is assigned to constituent N3 in the R/N2 compound. The comparison of the two pitch tracks illustrates that the crucial difference between L/N1 and R/N2 compounds is that of the pitch value assigned to constituent N2. In particular, according to the Embedded Prominence Hypothesis, we expect speakers to assign on average a higher pitch to

³The two pitch tracks are extracted from recordings of two different speakers. Because of that the pitch range displayed on the vertical axis of the two pitch tracks differs between the two tracks. Yet, the difference in the pitch range between the two speakers is accounted for by modelling *SPEAKER* as a random effect in a mixed-effects model (see again section 3.4 for more details on random effects in a mixed-effects model).

Figure 6.1: Pitch track of L/N1 compound: [*háy fever*] treatment

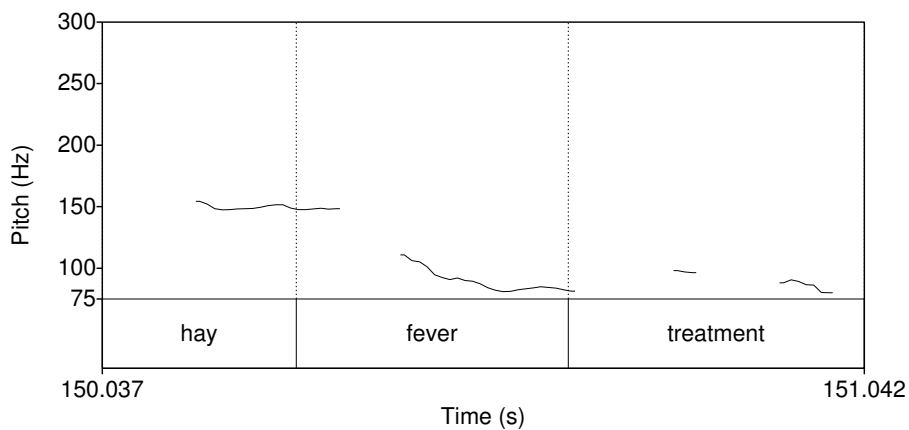
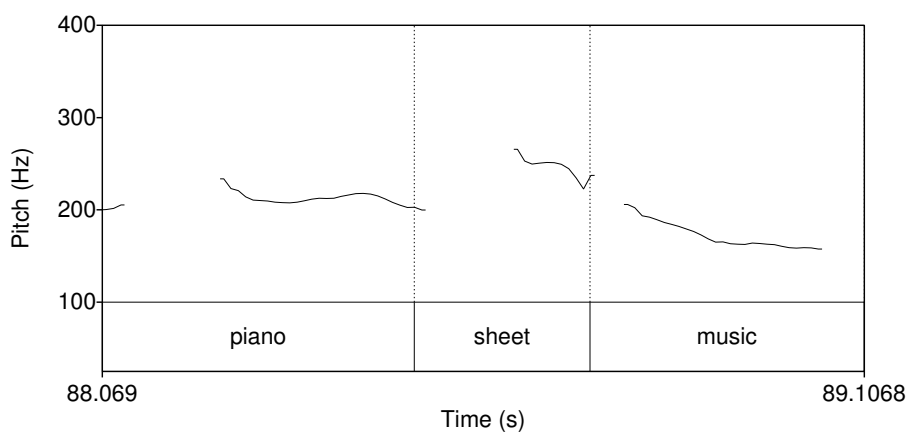


Figure 6.2: Pitch track of R/N2 compound: *piano* [*shéet music*]



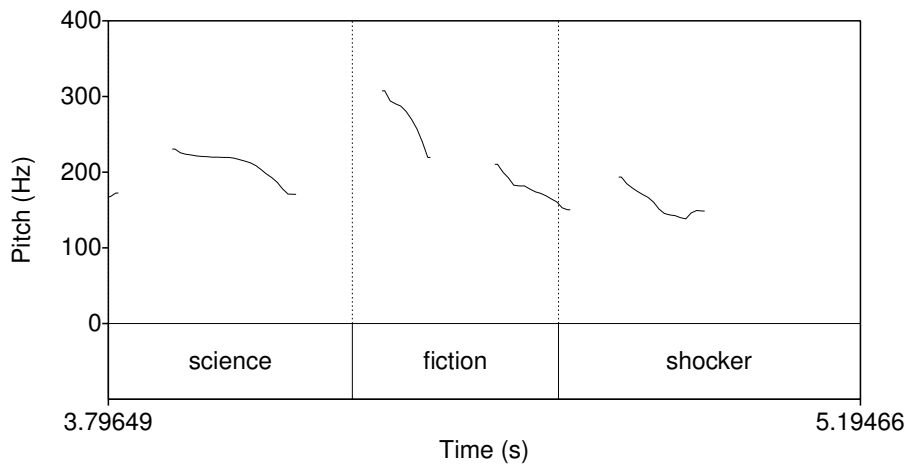
constituent N2 in right-branching compounds with embedded left prominent NNs than in left-branching compounds with embedded left prominent NNs. In contrast, with reference to the pitch values assigned to constituent N3, we do not expect a difference between the two compound groups, as in both, constituent N3 should be the least prominent constituent. Finally, we expect pitch on constituent N1 to be equally high in the two compound groups.

The second part of the Embedded Prominence Hypothesis makes predictions for left- and right-branching compounds with embedded right prominent NN compounds. According to the EPH, highest prominence is assigned to constituent N2 in a left-branching compound with an embedded right prominent NN and to constituent N3 in a right-branching compound whose complex constituent is right prominent. As described in subsection 6.1.1, the two compound groups created to test this part of the Embedded Prominence Hypothesis are that of L/N2 and R/N3 compounds. If the EPH is correct in its predictions, L/N2 and R/N3 compounds should differ in pitch from left- and right-branching compounds with embedded left prominent NNs, i.e. L/N1 and R/N2 compounds. Thus, the pitch values of L/N2 compounds had to be compared to those of L/N1 compounds, whereas the pitch values of R/N3 compounds had to be compared to the ones measured for R/N2 compounds.

Starting off with L/N2 compounds, a relatively high pitch is expected to be assigned to constituents N1 and N2, which would indicate a higher prominence of the second constituent in relation to the first; constituent N3 should have the lowest pitch in L/N2 compounds. In contrast, L/N1 compounds are only expected to exhibit a high pitch on constituent N1 with a clearly lower pitch on constituents N2 and N3. For a better illustration of the predicted pitch patterns, a pitch track of an L/N2 compound is provided in Figure 6.3. The pitch curve of an L/N1 compound is given in Figure 6.1.

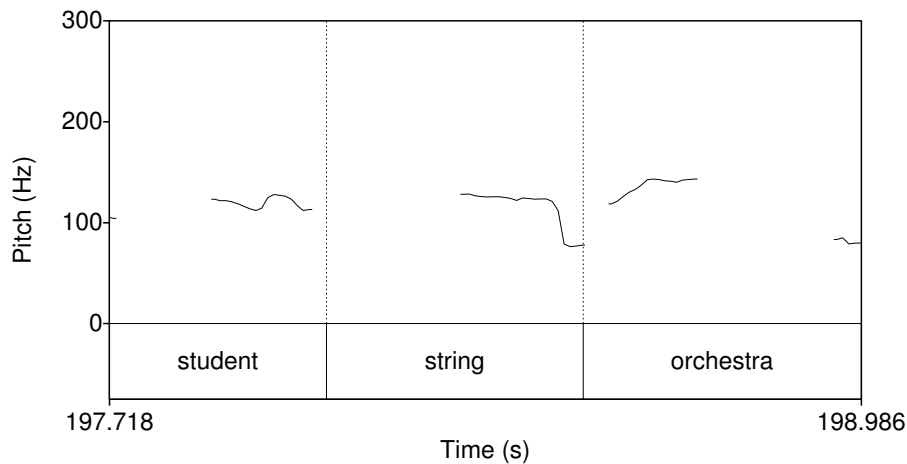
A comparison of the two pitch tracks clearly illustrates that we would expect a difference between the two compounds with reference to pitch assigned to constituent N2 by the speakers; pitch on constituent N2 should be higher in L/N2 compounds than in L/N1 compounds. Furthermore, L/N2 compounds may either have a lower pitch on constituent N1 than L/N1 compounds or an equally high pitch. The pitch values on constituent N3 are not expected to differ between the two groups. Hence, the predicted difference in the pitch pattern between L/N1 and L/N2 compounds is similar to that predicted for L/N1 and R/N2 compounds.

Finally, according to the Embedded Prominence Hypothesis, the group of R/N3 compounds should exhibit highest prominence on constituent N3 in contrast to R/N2

Figure 6.3: Pitch track of L/N2 compound: [*science fiction*] *shocker*

compounds in which constituent N2 should be the most prominent constituent. As mentioned above, R/N2 compounds should exhibit relatively high pitch values on both constituent N1 and N2 with a clearly lower pitch on constituent N3. In R/N3 compounds, highest prominence should be assigned to constituent N3. Therefore, we would expect R/N3 to have a higher pitch on constituent N3 than R/N2 compounds. Moreover, the pitch values on constituent N1 and constituent N2 are not expected to differ between the two compound type, as right prominence is phonetically marked by a high pitch on two consecutive constituents (e.g. Kunter 2011). Given this finding, we may expect speakers to use a relatively high pitch on all three constituents of the R/N3 compounds in order to mark constituent N3 to be the most prominent one of the compound.

Figure 6.4 shows a pitch track of an R/N3 compound for a better illustration of the predicted pitch pattern for this group. The example is again taken from the data set of the present experiment. The pitch track of the R/N3 compound *student string orchestra* shows high pitches on both constituent N1 and constituent N2. Furthermore, pitch slightly increases over the utterance of the third constituent so that constituent N3 has the highest pitch in this example. The pitch track of the R/N2 compound *piano sheet music* as given in Figure 6.2 differs from the one in Figure 6.4 in that pitch decreases on constituent N3 in the R/N2 compound.

Figure 6.4: Pitch track of R/N3 compound: *student* [string orchestra]

6.1.3 Statistical procedure

This section provides some details on the statistical modelling of the data. In a first step, i.e. before a mixed-effects model was fitted to the data (see subsection 3.4 again for more details on mixed-effects models), the distribution of the measured pitch values was inspected to check for the presence of extreme data points in the data set. The respective plots revealed indeed some striking values among the data. In order to find out whether those values could be restricted to either male or female speakers, the distribution of the pitch values for the two genders were inspected separately. The separate inspection of the two speaker subsets revealed for each of them a small number of pitch values which clearly stood out from the rest; these values were much lower and higher, respectively, than the majority of the remaining values (male ($N = 20$); female speakers ($N = 7$)). In order to avoid these pitch values to become overly influential in the subsequent statistical analysis, they were excluded from the data set prior to the statistical modelling of the data (cf. Baayen 2008:243f). This reduced the total number of observations from 1479 to 1452.

A linear mixed-effects model was fitted to the data with SPEAKER and ITEM as random effects, BRANCHING, POSITION and GENDER as fixed effects. The dependent variable was pitch in semitones (pitchST). The random effect SPEAKER was included in the model to account for between-speaker differences as for instance introduced by differences in the speakers' pitch range, the speakers' physical condition during the experiment or dialectal differences (see again section 3.4 for a more detailed discussion of between-speaker differences). Furthermore, ITEM was modelled as a random

effect in the model in order to account, among other things, for differences in the syllable structure of the test items or intrinsic pitch differences of the nuclei. By means of two likelihood ratio tests it was tested whether the inclusion of SPEAKER and ITEM as random effects in the model was justified (cf. Baayen 2008:253). In such likelihood ratio tests a model containing the random effect is compared to a simpler model without the respective factor fitted as a random effect. The two tests revealed a significant increase of the log likelihood in the two models that contained the random effects in contrast to the models from which they were removed. This indicated that the inclusion of SPEAKER and ITEM as random effects in the model was fully justified (the outcome of the two likelihood ratio tests are documented in part B of the Appendix).

The first fixed factor BRANCHING served as grouping factor and consisted of the four factor levels *ln1*, *ln2*, *rn2* and *rn3*. The levels represent the four constructed compound groups.⁴ The factor POSITION consisted of the three factor levels *lpitch*, *mpitch* and *rpitch* which represent the position in the compound for which pitch was measured.⁵ The third factor GENDER consisted of the two factor levels *male* and *female*. The factor was added to the model in order to account for differences in the pitch range between male and female speakers. As noted before, the pitch range for male speakers of English lies on average between 70Hz - 250Hz, whereas that of an adult female speaker roughly ranges between 80Hz - 400Hz (e.g. Gut 2009:172). In the light of these general pitch differences between the two genders, the factor GENDER was expected to be highly significant. Finally, a three-way interaction between the factors BRANCHING, POSITION and GENDER was added to the model to test whether all three factors also interacted with each other.⁶ Yet, the three-way interaction was not expected to be significant as there was no theoretically motivated reason to assume that male and female speakers would differ systematically in prominence assignment to compounds. It was rather added to the model to additionally explore the possibility of such an effect.

In contrast, the inclusion of the subordinate two-way interaction of BRANCHING by POSITION in the model was crucial for the present analysis. The interaction tells us whether the speakers assign different pitch patterns to the four compound groups. In

⁴It must be noted here that the factor name BRANCHING does not only refer to the branching direction of the compound but also encodes the crucial information about the prominence pattern of the complex constituent of the compound. See again section 6.1.1 for an explanation of the four abbreviations.

⁵The abbreviations 'lpitch', 'mpitch' and 'rpitch' stand for left pitch (N1), middle pitch (N2) and right pitch (N3).

⁶The interaction term for the three-way interaction was formulated in a way that all subordinate two-way interactions were tested for significance as well.

particular, it provides information on whether or not the four compound groups differ in the degree to which pitch drops between any of the three constituents. Such differences in the pitch drop are predicted by the Embedded Prominence Hypothesis. For instance, it is predicted that R/N2 compounds and L/N1 compounds exhibit equally high pitches on constituent N1 but differ significantly in pitch on constituent N2. This implies that pitch must drop to a different degree between the two constituents of the two compound groups, which is exactly the information provided by a significant interaction of BRANCHING by POSITION. Thus, in the absence of a significant interaction of BRANCHING by POSITION it would be implied that all compounds have the same pitch pattern. This again would suggest that the assignment of pitch, i.e. the assignment of prominence, would not depend on differences in the branching direction of the compounds nor on the prominence pattern of the embedded NN compounds, but that other factors must be at work.

After I fitted a first model to the data, the residuals of the model were checked for potential outliers. All data points with residuals greater a standard deviation than $|2.5|$ were excluded from the data set (e.g. Baayen 2008; Baayen and Milin 2010). This resulted in a loss of 54 (3.7%) observations, which reduced the total number of data points from 1452 to 1398. Interestingly, a closer inspection of the diagnostic plots of the first model revealed that 37 of the 54 removed observations belonged to one particular speaker in the data set. Figure 6.5 shows one of those diagnostic plots for a better illustration, with the residuals for that one particular speaker being marked by the crosses in the plot.⁷ Hence, the inaccuracy of the first model seemed to be primarily due to the performance of this one particular speaker.

A new model with the same model specifications as the previous one was refitted to the trimmed data set. The inspection of the residuals of the new model revealed that the previously found non-normality of the residuals had been largely removed. Yet, the new model still failed to predict a number of observations with high residuals in the right tail of the distribution. A closer inspection of these values showed that they belonged again primarily to the data set of this one particular speaker (see Figure 6.6).

Since the model had trouble to predict the pitch values of this one particular speaker, I decided to exclude all data points of that speaker from the data set in order to avoid his performance to become overly influential in the subsequent pitch analysis. Yet, I decided to additionally investigate the data of the excluded speaker in more detail in a separate analysis. The reason to do so was that all recordings had been previously checked for recording quality, natural reading and non-modal phonation. Hence, the

⁷The 37 observations made up one third of the speaker's total amount of measurements.

Figure 6.5: Standardized residuals of mixed-effects model before the removal of values with residuals exceeding a standard deviation of $|2.5|$.

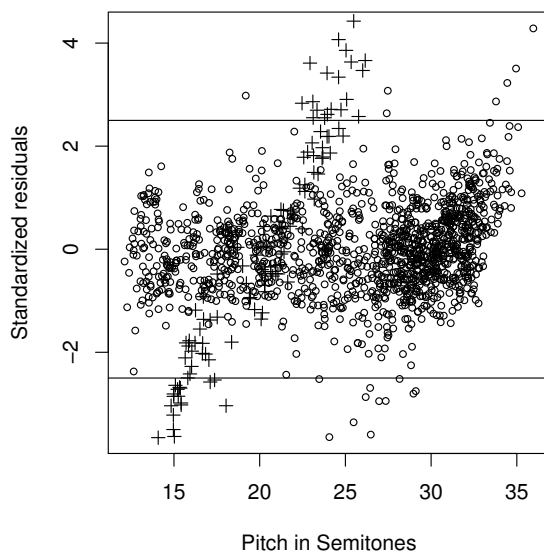
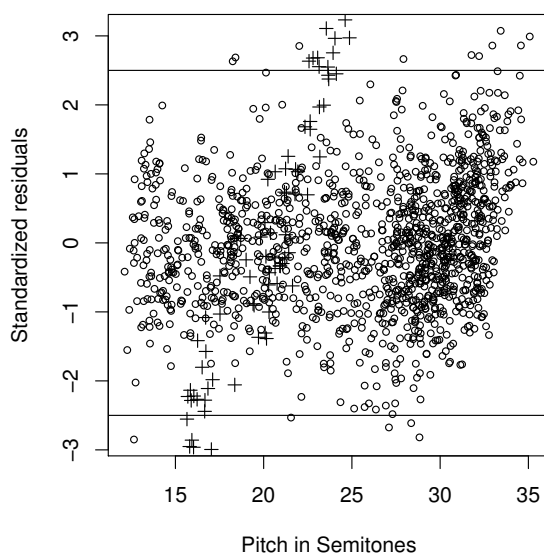


Figure 6.6: Standardized residuals of trimmed model, i.e. after the removal of values with residuals exceeding a standard deviation of $|2.5|$.



model's inaccuracy must have been caused by the speakers's idiosyncratic use of pitch rather than by factors such as reading performance or creaky voice. For instance, this speaker might have had constantly assigned a different pitch pattern to one of the four compound groups than the other 12 speakers which is why the model failed to predict the speaker's pitch values. Another possibility might have been that the excluded speaker varied pitch extremely within his own pitch range which resulted in extremely low and high pitch values. Therefore, in order to gain a better understanding of the speaker's performance, I also had a look at his data set. The separate investigation of the speaker's data set is described in section 6.4.

The removal of the speaker's data points reduced the total number of observations from 1452 to 1340. A new model was fitted to the trimmed data set from which this speaker's data had been removed. Following the standard procedure of model criticism (cf. Baayen 2008), observations with residuals exceeding a standard deviation of $|2.5|$ were excluded from the data set. This concerned 32 observations (2.8%). The inspection of the respective diagnostic plot of the trimmed model showed that the non-normality of the residuals was largely removed. The total number of observations was reduced from 1340 to 1308.

During subsequent model simplification, non-significant parameters were removed from the model in a step-wise fashion beginning with the highest-order interaction (cf. Crawley 2005b:105). In the current model this was the three-way interaction of BRANCHING by POSITION by GENDER. By means of model comparison the model containing the three-way interaction was compared to a simpler model from which the interaction parameter was removed. The comparison revealed that the simpler model was not significantly worse in its explanatory power than the more complex model, which justified the exclusion of the three-way interaction (cf. Crawley 2005a:262). Hence, there was no evidence in the data that male and female speakers produced different pitch patterns for the four compound groups.

Similar tests for all subordinate two-way interactions resulted in an exclusion of the interaction GENDER by BRANCHING.⁸ Yet, the two-way interactions of GENDER by POSITION and BRANCHING by POSITION remained significant despite the removal of the other interaction terms. The significant interaction of BRANCHING by POSITION indicates that at least two compound groups differ from each other in the average pitch pattern assigned to them. The interaction of GENDER by POSITION indicates

⁸The order in which the two-way interactions were removed from the model during model simplification was determined by the p -values obtained in an ANOVA of the model; the least significant term was removed first (cf. eg. Crawley 2005b:105).

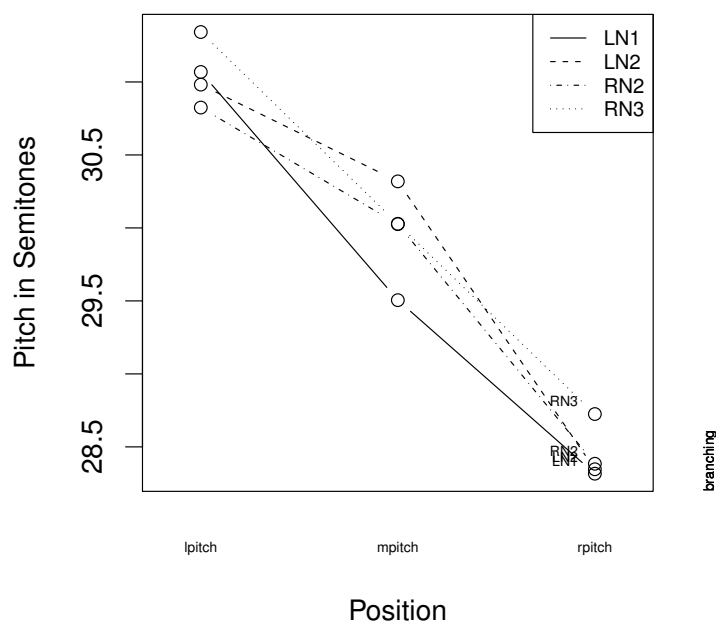
that males and females in general slightly differ in the degree to which pitch drops between two constituents. This effect, however, is less relevant for the present analysis than the one of BRANCHING by POSITION.

Finally, according to the Embedded Prominence Hypothesis (see section 6.1.2 again) L/N2 and R/N2 compounds should both exhibit highest prominence on constituent N2, irrespective of their different internal structures. This similarity in the prominence behaviour of the two groups should manifest itself in similar pitch patterns of L/N2 and R/N2 compounds. Because of that, I checked whether it was possible to combine the two factor levels *rn2* and *ln2*, which represent the two compound groups, into a single factor level in the model. If the combination of the two factor levels is possible without reducing the explanatory power of the statistical model, we know that L/N2 and R/N2 compounds are statistically the same. This provides first evidence that the prediction of the EPH with regard to these two compounds groups may be in fact correct. Hence, in order to find out if the combination was possible, the current model was compared to a new model in which the factor levels *rn2* and *ln2* were combined to a single factor level, namely *lrn2*. The comparison of the two models revealed that a combination of the two factor levels was indeed possible without reducing the explanatory power of the model (the outcome of the model comparison is documented in B of the Appendix). Hence the two groups have similar pitch patterns. Given this outcome, the model in which the two factor levels *rn2* and *ln2* were collapsed was preferred to the model in which they were treated as two separate levels. Hence, in the final mixed-effects model, the factor BRANCHING consisted only of three factor levels, namely *ln1*, *rn3* and *lrn2*.

For a better illustration of why it was possible to combine the two factor levels *rn2* and *ln2* to a single factor level, an interaction plot of the model in which the four compound groups are still treated as four separate factor levels is given in Figure 6.7. In the plot, pitch in semitones is given on the vertical axis, while the horizontal axis shows the three different positions for which pitch was measured. The circles represent the mean values predicted by the model at each position in each compound. The different compound groups are represented by the different lines, which connect the mean pitch values of the three positions in each compound. L/N1 compounds are represented by the solid line, L/N2 compounds by the dashed line, R/N2 compounds by the dotted line, and R/N3 compounds by the dotted-dashed line.

The plot shows that the pitch values of constituent N1 and constituent N2 for L/N2 compounds are slightly higher than that of the R/N2 compounds. Yet, pitch drops almost to the same degree between each constituent in the two groups, which is indi-

Figure 6.7: Interaction plot displaying the average mean pitches estimated by the mixed-effects model for each constituent of L/N1, L/N2, R/N2 and R/N3 compounds; female speakers.



cated by the parallel lines connecting the mean values. Having discussed all preliminary statistical measures, I will now turn to the result of the statistical analysis.

6.2 Results

The aim of the present experiment is to test whether in left- and right-branching compounds with embedded left prominent NNs highest prominence is assigned to constituent N1 and constituent N2, respectively. Furthermore, it is tested whether the embedding of a right prominent NN compound in a left- and right-branching compound causes highest prominence on constituent N2 in left-branching compounds and on constituent N3 in right-branching compounds. This chapter presents the results of the statistical analysis. Section 6.2.1 provides the result of the pitch analysis for left- and right-branching compounds with embedded left prominent NNs. Section 6.2.2 presents the result for left- and right-branching compounds with embedded right prominent NNs. As in the corpus study presented in chapter 4, all p -values presented for the fixed effects coefficients of the mixed-effects model in this chapter (see Table 6.3 below) were obtained by means of Markov chain Monte Carlo sampling

(MCMC) (cf. Baayen 2008:248). Furthermore, treatment contrasts were used to obtain all contrasts relevant to test the hypothesis at hand.⁹

6.2.1 Embedded Prominence Hypothesis: Embedded left prominent NN compounds

The first comparison to be investigated was that between the pitch values of L/N1 compounds and R/N2 compounds. According to the Embedded Prominence Hypothesis, the two compound groups should differ from each other in that R/N2 compounds exhibit a higher pitch on constituent N2 than L/N1 compounds. The pitch values of constituent N1 and N3 were not expected to differ between the two groups.

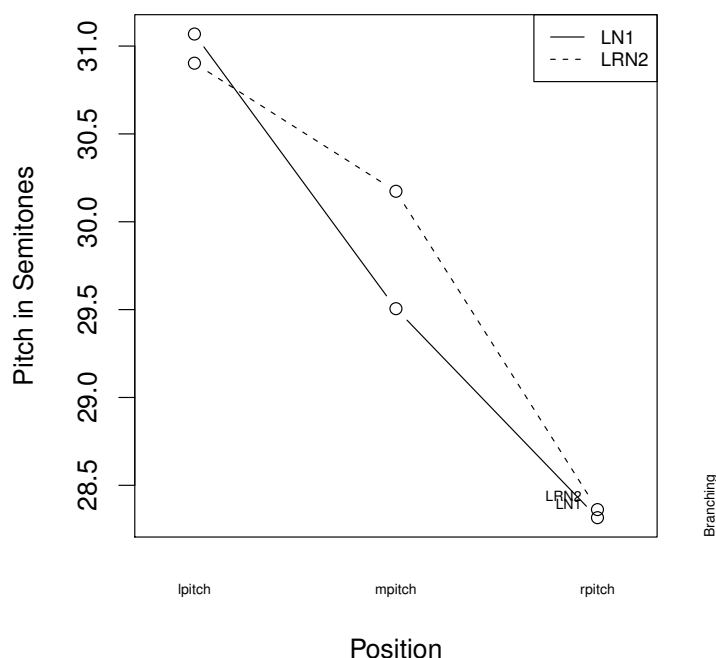
In order to gain a first visual impression of the result, Figure 6.8 shows an interaction plot, which displays the mean pitch values for R/N2 and L/N1 compounds as predicted by the model for female speakers. The corresponding pitch values for male speakers are about 10ST lower for each position in R/N2 and L/N1 compounds (for more details see Table 6.3). Pitch in semitones is given on the vertical axis and the three relevant pitch positions of the compounds are given on the horizontal axis. The R/N2 compounds are represented by the dashed line, which is labelled LRN2, and the L/N1 compounds by the solid line.¹⁰ The mean pitch values estimated at each position in the two compound groups are marked by the circles.

The plot shows that R/N2 and L/N1 compounds start off with roughly the same amount of pitch on constituent N1 (*lpitch*). Yet, we observe that pitch drops to a larger degree between constituent N1 and N2 in L/N1 compounds (1.56ST) than in R/N2 compounds (0.73 ST), which results in a clearly lower pitch on constituent N2 in L/N1 compounds than in R/N2 compounds. This difference in pitch on constituent N2 is indicated by the large gap between the two circles representing the mean pitch at position *mpitch* in R/N2 compounds (30.2 ST) and L/N1 compounds (29.5 ST). Finally, the plot shows that R/N2 and L/N1 compounds exhibit almost equally high pitch values (about 28.3 ST) on constituent N3 (*rpitch*), which is due to a steeper pitch fall between constituent N2 and N3 in R/N2 compounds (1.81ST) as compared to that in L/N1 compounds (1.19 ST). The interaction plot strongly suggests that R/N2

⁹The R-programme uses treatment contrasts as its default convention. Thus, the factor level that comes alphabetically and numerically first is automatically chosen as the model's baseline. In the present data set the baseline was *lpitch*, *female* and *ln1*. If necessary, this baseline was altered by changing the order of the factor levels of BRANCHING, POSITION and GENDER, respectively.

¹⁰The line representing the group of R/N2 compounds is labelled 'LRN2' due to the fact that the two factor levels *rn2* and *ln2* were collapsed into the single factor level *lrn2* (see again section 6.1.3 on that matter).

Figure 6.8: Interaction plot displaying the average mean pitches estimated by the mixed-effects model for each constituent of L/N1 and R/N2 compounds; female speakers.



compounds differ from L/N1 compounds according to the Embedded Prominence Hypothesis, namely in pitch assigned to constituent N2. In order to know whether the observed differences are also statistically significant, however, it is necessary to look at the actual outcome of the statistical analysis.

The baseline in Table 6.3 is that of `ln1`, `mpitch` and `female`, i.e. the estimated mean pitch of constituent N2 in L/N1 compounds of female speakers (29.5 ST). The corresponding N2 mean pitch of L/N1 compounds for male speakers may be obtained by subtracting the coefficient of `male` from the intercept of the model. As mentioned in the beginning of this section, the negative coefficient of `male` tells us that the estimated mean pitch values for constituent N2 of the male speakers are generally 10.37 ST lower than that of the female speakers. The two interaction coefficients of GENDER by POSITION indicate that male speakers use in general a slightly higher pitch on constituent N1 and constituent N3 than female speakers. In order to obtain these estimated mean pitches for constituent N1 and N3 of the male speakers, the interaction coefficients of GENDER by POSITION must be added to the coefficient of `male`.¹¹

¹¹As mentioned in section 6.1.3, the interaction of GENDER by POSITION is not crucial for the present analysis since it does not capture the information that we are dealing with three different compound

Table 6.3: Table of coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: BRANCHING = ln1, POSITION = mpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	29.5053	1.1897	24.801	< 0.001
ln2	0.6681	0.2056	3.250	< 0.01
rn3	0.5232	0.2379	2.200	< 0.05
lpitch	1.5632	0.1341	11.658	< 0.001
rpitch	-1.1896	0.1371	-8.680	0.001
male	-10.3700	1.8259	-5.679	0.001
ln2:lpitch	-0.8341	0.1514	-5.511	< 0.001
rn3:lpitch	-0.2498	0.1746	-1.431	0.153
ln2:rpitch	-0.6220	0.1541	-4.036	< 0.001
rn3:rpitch	-0.1146	0.1775	-0.645	0.519
lpitch:male	0.1124	0.1257	0.895	0.371
rpitch:male	0.3724	0.1280	2.909	<0.01

Turning to the relevant pairwise comparisons between R/N2 and L/N1 compounds, the table shows that the interaction the two interaction coefficients of `ln2:lpitch` (-0.8431) and `ln2:rpitch` (-0.6221) are highly significant ($p < 0.001$). Hence, the pitch drop between constituent N1 and N2 is indeed statistically significantly smaller in R/N2 compounds than in L/N1 compounds (see Figure 6.8 again). Furthermore, pitch drops to a statistically significantly larger extent between constituent N2 and N3 in R/N2 compounds than in L/N1 compounds. Most importantly, however, the table also shows that the relevant contrast between the estimated mean pitches for constituent N2 of L/N1 and R/N2 compounds (`ln2`) is statistically significant at the 0.5 significance level ($p < 0.01$). Thus, the model provides evidence that speakers assigned systematically higher pitches to constituent N2 in R/N2 compounds than in L/N1 compounds, a result fully in accordance with the Embedded Prominence Hypothesis.

In addition to the contrasts displayed in Table 6.3, I also investigated whether R/N2 and L/N1 compounds differed from each other in the estimated pitches for constituent N1 and constituent N3, respectively.¹² When discussing the interaction plot given in Figure 6.8, it has already been noted that R/N2 and L/N1 compounds had al-

groups. The interaction parameter, however, was kept in the model for sake of completeness. Yet it will not be discussed in any more detail here.

¹²In order to obtain these contrasts, the baseline of the model had to be changed from `mpitch` to `lpitch` and `rpitch`, respectively.

most equally high pitch values on constituent N1 and on constituent N3. In fact, this visual impression is supported by the statistical analysis. The coefficient capturing the difference between the mean pitch values of constituent N1 was found to be non-significant ($p = 0.421$) (see Table 6.5 in section 6.2.2). A similar result was obtained for the comparison of the two pitch means of constituent N3 of R/N2 and L/N1 compounds; the p -value of the respective coefficient was found to be non-significant at the 0.5 significance level ($p = 0.825$) (see Table 6.4 given in section 6.2.2).

The result of the statistical analysis provides evidence that the pitch patterns of L/N1 and R/N2 compounds are in accordance with the Embedded Prominence Hypothesis. R/N2 compounds exhibit a higher pitch on constituent N2 than L/N1 compounds. Furthermore, R/N2 and L/N1 compounds do not significantly differ from each other with reference to the mean pitches assigned to constituent N1 and N3.

6.2.2 Embedded Prominence Hypothesis: Embedded right prominent NN compounds

The previous subsection presented the result of the pitch analysis for left- and right-branching compounds with embedded left prominent NNs. In this subsection the result of the pitch analysis for left- and right-branching compounds with embedded right prominent NNs is presented. We begin this subsection with the result for the group of left-branching compounds with embedded right prominent NNs (L/N2) before turning to the result for the group of right-branching compounds with embedded right prominent NNs (R/N3). According to the Embedded Prominence Hypothesis, L/N2 compounds exhibit highest prominence on constituent N2. Hence, the Embedded Prominence Hypothesis makes the same prominence prediction for the group of L/N2 compounds as it does for the group of R/N2 compounds. Because of that, we would also expect the same pitch differences between the three members of the L/N1 and L/N2 compounds as those observed between L/N1 and R/N2 compounds. In particular, we would expect left-branching compounds with embedded right prominent NNs to exhibit a higher pitch on constituent N2 than left-branching compounds with embedded left prominent NNs. With reference to the pitch values assigned to constituent N1 and N3, however, the two left-branching compound groups are not expected to differ from each other (see section 6.1.2 again for details on this prediction).

The statistical modelling of the data has already revealed (see subsection 6.1.3 again) that the pitch patterns of L/N2 and R/N2 compounds were indeed similar in the present data set. This was indicated by the fact that it was possible to collapse the

two factor levels *rn2* and *ln2*, which represent the R/N2 and L/N2 compounds, respectively, into a single factor level without reducing the model's explanatory power. Hence, L/N2 and R/N2 compounds are both represented by the factor level *l_rn2* in the final model. As a consequence, the result of the pitch comparison between left- and right-branching compounds with embedded left prominent NNs as presented in the previous subsection is at the same time the result of the comparison of L/N1 and L/N2 compounds. Hence, L/N2 compounds exhibit a significantly higher pitch on constituent N2 than L/N1 compounds (see again the coefficient of *l_rn2* in Table 6.3 as well as Figure 6.8 for a visual illustration of the result). Furthermore, the two left-branching compound groups neither differ from each other regarding the mean pitches estimated for constituent N1 nor with respect to those returned by the model for constituent N3. Thus, L/N2 compounds differ from L/N1 compounds only in pitch assigned to constituent N2, which is in line with the prediction of the Embedded Prominence Hypothesis.

This leaves us with the question whether the Embedded Prominence Hypothesis also predicts the correct prominence pattern for right-branching compounds with embedded right prominent NNs. According to the EPH, right-branching compounds with a right prominent complex constituent should have highest prominence on constituent N3, whereas right-branching compounds with a left prominent complex constituent should have highest prominence on constituent N2. With reference to the pitch comparison of the two right-branching compound groups, we would expect R/N3 compounds to exhibit a statistically significantly higher pitch on constituent N3 than R/N2 compounds. In contrast, with respect to the pitch values assigned to constituent N1 and constituent N2, we would not expect a difference between the two compound groups.

Table 6.4 shows the table of coefficients of the model with the baseline representing the estimated mean pitch of constituent N3 (*rpitch*) in R/N2 compounds (*l_rn2*) for female speakers (*female*).¹³ Figure 6.9 shows an interaction plot of the model that displays the two groups of R/N2 and R/N3 compounds. Pitch in semitones is given on the vertical axis and the three pitch positions on the horizontal axis. R/N3 compounds are represented by the dotted line, R/N2 compounds by the solid line. The estimated mean pitch values at each position of R/N2 and R/N3 compounds are marked by circles.

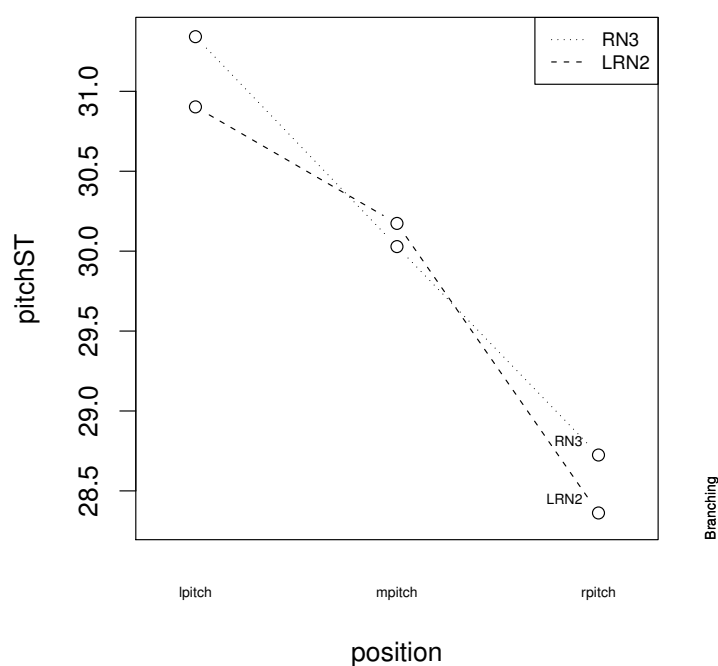
Figure 6.9 shows that pitch drops in both R/N2 and R/N3 compounds over the ut-

¹³An alternation of the contrast coding was used to change the baseline of the model for the factors POSITION and BRANCHING.

Table 6.4: Table of coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: BRANCHING = lrn2, POSITION = rpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	28.3619	1.1841	23.952	< 0.001
ln1	-0.0461	0.2082	-0.221	0.825
rn3	0.3626	0.2083	1.741	0.082
lpitch	2.5407	0.1056	24.060	< 0.001
mpitch	1.8116	0.1054	17.189	< 0.001
male	-9.9976	1.8261	-5.475	0.001
ln1:lpitch	0.2121	0.1551	1.368	0.172
rn3:lpitch	0.0768	0.1541	0.498	0.619
ln1:mpitch	-0.6220	0.1541	-4.036	< 0.001
rn3:mpitch	-0.5075	0.1490	-3.276	< 0.01
lpitch:male	-0.2600	0.1287	-2.020	< 0.05
mpitch:male	-0.3724	0.1280	-2.909	< 0.01

Figure 6.9: Interaction plot displaying the average mean pitches estimated by the mixed-effects model for each constituent of R/N3 and R/N2 compounds; female speakers.



terance of the compounds. As mentioned in previous subsections, we may interpret this general pitch drop as being the result of a general pitch declination effect, which is due to a natural decline of vocal cord vibrations during the course of an utterance. Yet, what is more important for the present analysis than this general pitch drop is the observation that the R/N2 and R/N3 compounds differ in the degree to which pitch drops between each of the three constituents. This strongly suggests that the R/N2 and R/N3 compounds are stressed differently by the speakers. In particular, the plot shows that the mean pitch of constituent N1 in R/N3 compounds is about 0.5 semitones higher than the mean pitch of R/N2 compounds. Furthermore, the mean pitch of constituent N2 of R/N3 compounds (30.03 ST) tends to be slightly lower than that of R/N2 compounds (30.17 ST). According to the plot, this lower pitch on constituent N2 in R/N3 compounds is due to a larger pitch drop between constituent N1 and constituent N2 in R/N3 compounds than in R/N2 compounds. More important, however, with reference to the Embedded Prominence Hypothesis is the observation that pitch drops less steeply between constituent N2 and constituent N3 in R/N3 compounds than it does in R/N2 compounds. Crucially, this smaller pitch drop in R/N3 compounds results in a higher mean pitch on constituent N3 in R/N3 compounds than in R/N2 compounds, which is also highly in accordance with the EPH.

A look at Table 6.4 shows that the observed difference in the pitch drop between constituent N2 and constituent N3 of R/N2 and R/N3 compounds is also statistically significant ($\text{rn3:mpitch} = p < 0.01$). However, contrary to the prediction of the EPH, the significantly smaller pitch drop between constituent N2 and constituent N3 in R/N3 compounds does not result in a statistically significantly higher mean pitch of constituent N3 in R/N3 compounds. This is indicated by the non-significant coefficient of rn3 ($p = 0.082$), which represents the contrast between the mean pitches estimated for constituent N3 of R/N3 and R/N2 compounds. The result indicates that the pitch values obtained for constituent N3 in the two compound groups are not different enough in order for the mean values to be statistically significantly different from each other. Hence, the result provides no statistical support for the Embedded Prominence Hypothesis and its prediction that right-branching compounds with embedded right prominent NNs have a generally higher pitch, i.e. higher prominence, assigned to constituent N3 than R/N2 compounds.

Given the fact that the predicted effect for R/N3 compounds is absent, the question arises of how R/N3 compounds are actually stressed by the speakers. This question is in so far of interest as we noted above that R/N2 and R/N3 compounds slightly differ in pitch assigned to constituent N1 as well as in the degree to which pitch drops be-

tween constituent N1 and N2. This may suggest that R/N3 compounds, after all, exhibit a different prominence pattern than R/N2 compounds. In particular, the higher pitch on constituent N1 and the lower pitch on constituent N2 suggests that in R/N3 compounds highest prominence tends to be assigned to constituent N1, which would be a violation of the Embedded Prominence Hypothesis at the IC-level of the compounds. In order to find out whether these differences observed in Figure 6.9 are statistically significant, I also investigated the remaining contrasts between R/N2 and R/N3 compounds.

Table 6.5 shows the table of coefficients of the final model, yet this time with the baseline representing the mean pitch estimated for constituent N1 in R/N2 compounds for female speakers. The table shows that the mean pitch of constituent N1 in R/N3 compounds is 0.43 semitones higher than in R/N2 compounds. This difference is statistically significant according to the p -value of the respective coefficient ($p < 0.05$). In addition to that, we note that the interaction coefficient capturing the difference in the pitch drop between constituent N1 and constituent N2 (rn3:mpitch) is highly significant ($p < 0.01$). Thus, apart from the higher pitch on constituent N1, we also find that the pitch drop between constituent N1 and constituent N2 is significantly larger in R/N3 compounds than in R/N2 compounds. Finally, the coefficient capturing the contrast between the mean pitch values estimated for constituent N2 in R/N2 and R/N3 compounds shows that the difference is non-significant according to the model.¹⁴ Hence, R/N2 and R/N3 compounds do not differ significantly in pitch on constituent N2.

¹⁴The table of coefficients that displays the contrast between the mean pitches of constituent N2 of R/N2 and R/N3 compounds is given in part B of the Appendix.

Table 6.5: Table of coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: BRANCHING = lrn2, POSITION = lpitch, GENDER = female

	Estimate	Std. Error	t-value	p - value
(Intercept)	30.9026	1.1838	26.105	< 0.001
ln1	0.1659	0.2063	0.804	0.421
rn3	0.4394	0.2056	2.137	< 0.05
mpitch	-0.7291	0.1016	-7.175	< 0.001
rpitch	-2.5407	0.1056	-24.060	< 0.001
male	-10.2575	1.8269	-5.618	< 0.001
ln1:mpitch	-0.8341	0.1514	-5.511	0.001
rn3:mpitch	-0.5843	0.1514	-3.860	0.001
ln1:rpitch	-0.2121	0.1551	-1.386	0.619
rn3:rpitch	-0.0768	0.1542	-0.498	0.371
mpitch:male	-0.1124	0.1257	-0.895	0.371
rpitch:male	0.2600	0.1287	2.020	0.044

Hence, the pitch comparison of R/N2 and R/N3 compounds revealed a statistically significant difference in the pitch patterns of the two right-branching compound groups. Yet, contrary to the Embedded Prominence Hypothesis, the R/N2 and R/N3 compounds do not differ in pitch assigned to constituent N3. Instead, we find that R/N3 compounds have a statistically significantly higher pitch on constituent N1 than R/N2 compounds. Furthermore, speakers drop their pitch on average to a larger degree between constituent N1 and N2 in R/N3 compounds than in R/N2 compounds. Yet, with reference to pitch assigned to constituent N2 and constituent N3, we find no difference between R/N2 and R/N3 compounds. The result is discussed in more detail in the following section.

6.3 Discussion

The production experiment presented in this section tested the predictions of the Embedded Prominence Hypothesis for left- and right-branching compounds with embedded left and right prominent NN compounds. The Embedded Prominence Hypothesis differs from Liberman and Prince's LCPR prediction in that it predicts prominence patterns of left- and right-branching compounds with embedded right prominent NN compounds. Such compounds are ignored by the LCPR given that right prominent NN compounds are thought to be phrases by Liberman and Prince

and as such should not occur embedded in triconstituent compounds. The Embedded Prominence Hypothesis predicts that if highest prominence falls on the complex constituent, left- and right-branching compounds with embedded left prominent NN compounds have highest prominence on the left member of the complex constituent whereas left- and right-branching compounds with embedded right prominent NN compounds are expected to exhibit highest prominence on the right member of the complex constituent.

In the present experiment, the prominence patterns of the compounds were determined by measuring the fundamental frequency, i.e. pitch, as an acoustic correlate to prominence. It has been shown (e.g. Kunter 2011; Plag et al. 2008) that a difference between left and right prominent compounds is marked by a difference in pitch assigned to the right member of the NN compound; in particular, right prominent NN compounds tend to have a higher pitch on N2 than left prominent NNs. Given this finding, the predicted difference in the prominence pattern of the four compound groups was expected to be marked by a difference in their pitch pattern. In particular, right-branching compounds with embedded left prominent NN compounds (R/N2) were expected to have a high pitch assigned to both constituent N1 and constituent N2, but a clearly lower pitch to constituent N3. A similar pitch pattern to that of R/N2 compounds was expected for the group of left-branching compounds with embedded right prominent NN compounds (L/N2) as they were also expected to exhibit highest prominence on constituent N2.

In contrast to that, left-branching compounds with embedded left prominent NN compounds (L/N1) were expected to have only a high pitch assigned to constituent N1; pitch on constituent N2 and on constituent N3 in L/N1 compounds was assumed to be clearly lower than the pitch on constituent N1. Thus, a comparison of the pitch values of L/N1, L/N2 and R/N2 compounds, was expected to reveal a significantly higher pitch on constituent N2 in R/N2 and L/N2 compounds than in L/N1 compounds. Finally, right-branching compounds with embedded right prominent NN compounds (R/N3) were expected to have relatively high pitches on all three compound constituents. Hence, when comparing the pitch values of R/N3 and R/N2 compounds, we expected to find a significantly higher pitch on constituent N3 in R/N3 compounds.

The statistical analysis of the data revealed that with the exception of right-branching compounds with embedded right prominent NN compounds (R/N3), the predictions of the Embedded Prominence Hypothesis are supported by the data. The pitch comparison of R/N2 and L/N1 compounds revealed a significantly higher pitch on

constituent N2 in R/N2 compounds. Crucially, at the same time speakers assigned on average equally high pitches to constituent N1 in R/N2 and L/N1 compounds as well as equally low pitches to constituent N3. Thus, the predicted higher prominence of constituent N2 in R/N2 compounds is indeed marked by a higher pitch on the respective constituent in these compounds. The result strongly suggests that in right-branching compounds with embedded left prominent NNs the most prominent constituent is in fact constituent N2, whereas in L/N1 compounds speakers marked constituent N1 as being the most prominent one.

In addition to that, left-branching compounds with embedded right prominent NN compounds (L/N2) were found to exhibit the same pitch pattern as right-branching compounds with embedded left prominent NN compounds (R/N2). Given that R/N2 compounds have highest prominence on constituent N2, we can claim the same for L/N2 compounds. This result provides strong evidence for the hypothesis that the embedding of a right prominent NN compound causes left-branching compounds to have highest prominence on the right member of the complex constituent, *contra* Liberman and Prince's (1977) LCPR prediction. Yet, the predicted effect for embedded right prominent NN compounds seems only supported for the group of left-branching compounds. The result for right-branching compounds with embedded right prominent NN compounds (R/N3) does not provide any evidence for this assumption.

The pitch analysis of R/N3 compounds revealed that, *contra* to the Embedded Prominence Hypothesis, the speakers did not assign a significantly higher pitch to constituent N3 in R/N3 compounds than in R/N2 compounds. Thus, there is no empirical support for the assumption that constituent N3 is more prominent in R/N3 compounds than in R/N2 compounds. This result raises two questions. First, how are R/N3 compounds stressed by the speakers if highest prominence is not assigned to constituent N3? Second, why is the Embedded Prominence Hypothesis supported for L/N2 compounds but not for R/N3 compounds?

With reference to the first question, an additional analysis of the pitch values of constituent N1 and constituent N2 showed that R/N3 compounds tended to have on average a significantly higher pitch on constituent N1 than R/N2 compounds. In addition to that, the pitch drop between constituent N1 and constituent N2 was significantly larger in R/N3 compounds. Yet, the pitch values assigned to constituent N2 in each group did not differ significantly from each other. Thus, according to the model the two right-branching compounds did not differ in pitch assigned to constituent N2 and constituent N3, but differed in pitch assigned to constituent N1. But what does this result tell us about the prominence pattern of the R/N3 compounds? The inter-

pretation of the result is somewhat problematic. The higher pitch on constituent N1 as well as the larger pitch drop between constituent N1 and constituent N2 in R/N3 compounds strongly suggest that speakers tended to assign highest prominence to constituent N1 in R/N3 compounds. Hence, there is a trend for R/N3 compounds to violate the Embedded Prominence Hypothesis at the IC-level of the compounds. However, what may speak against this interpretation of the result is the fact that R/N3 compounds differ from R/N2 compounds in pitch assigned to constituent N1 and not constituent N2. Thus far, compounds with highest prominence on constituent N1 (e.g. L/N1 compounds) differed from compounds with highest prominence on constituent N2 (e.g. R/N2 compounds) in that they exhibited a lower pitch on constituent N2. This is not the case, however, if we compare the R/N2 and R/N3 compounds. The fact that R/N3 compounds did not exhibit a lower pitch on constituent N2 than R/N2 compounds is therefore difficult to explain. Yet, it may be possible that among the R/N3 compounds we also find some compounds with highest prominence assigned to constituent N2, which may have caused the average pitch on that constituent to be relatively high. Hence, the speakers might have varied a lot in prominence assignment to these compounds. Because of this variation, an interpretation of the average pitch pattern in terms of prominence is difficult.

Let us turn to the second question raised above, namely why the Embedded Prominence Hypothesis is supported for L/N2 compounds but not for R/N3 compounds. In order to answer this question it is necessary to note again under which premises R/N3 compounds were expected to have highest prominence on constituent N3. First, it was expected that right prominence of the embedded constituent is preserved under embedding. Second, right-branching compounds had to be right prominent at their IC-level in order for the effect of the embedded NN to surface. Given these two assumptions, it may be possible that the predicted effect for the embedded right prominent NN is simply not supported by the given data because some factors independently operating at the IC-level of the compounds may have intervened and caused highest prominence on constituent N1 in some of the RN3 compounds.

Summarizing the discussion of the results, we found empirical evidence that in L/N1 and R/N2 compounds highest prominence is assigned to the left member of the complex constituent. Furthermore, the analysis revealed that in L/N2 compounds constituent N2 is the most prominent one, despite the fact that L/N2 compounds are left-branching. The result shows that the prominence pattern of an embedded right prominent NN is preserved under embedding and may cause the entire compound to have highest prominence on the right member of the complex constituent. How-

ever, for R/N3 compounds the predicted effect of an embedded right prominent NN is not supported by the data. As mentioned above, this might have been due to unknown factors operating independently at the IC-level of the compounds, which may have caused speakers to assign highest prominence to constituent N1 rather than to constituent N3 in some of the R/N3 compounds.

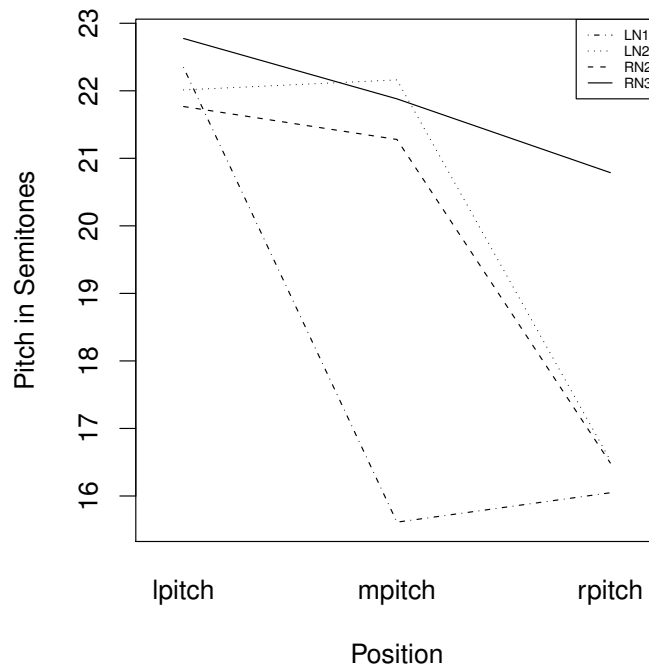
6.4 Individual speaker variation

It was mentioned in section 6.1.3 that the pitch measurements of one particular speaker were excluded during the statistical modelling of the data as it turned out that the model was particularly stressed by the speaker's observations. In this section the data of that speaker is explored in more detail in order to gain a better understanding of why the model had such trouble predicting the observations. In order to gain a first visual impression of the data, Figure 6.10 shows an interaction plot that displays the speaker's average mean pitches for each constituent in the four compound groups. Pitch in semitones is given on the vertical axis, whereas the horizontal axis displays the three positions for which pitch was measured in each compound. The dashed-dotted line in the plot represents the L/N1 compounds, the dashed line the R/N2 compounds, the dotted line the L/N2 compounds and the solid line the group of the R/N3 compounds.

Figure 6.10 shows that highest pitch is assigned to constituent N1 in all four compound groups. In particular, we observe that the average pitch on constituent N1 is about 22 semitones high in all groups, with only R/N3 compounds having a slightly higher pitch on constituent N1 than the other three compound types. Furthermore, we observe that pitch on constituent N2 (mpitch) is about 6 semitones lower in L/N1 compounds than in the other three compound groups. This is due to a pitch drop of about 6.5 semitones between constituent N1 and constituent N2. In contrast, pitch drops only 0.5 semitones between constituent N1 and constituent N2 in R/N2 compounds and 1 semitone in R/N3 compounds. Last but not least, for the group of L/N2 compounds, we even find a slight increase in pitch between constituent N1 and constituent N2 (+0.15ST).

Turning to the pitch values obtained for constituent N3 (rpitch), we observe that in L/N1 compounds pitch on constituent N3 is slightly higher than on constituent N2. Yet, pitch on constituent N3 in L/N1 compounds is still 6 semitones lower than that observed for constituent N1. In contrast, in L/N2, R/N2 and R/N3 compounds pitch decreases between constituent N2 and constituent N3. However, we observe a crucial

Figure 6.10: Interaction plot displaying the average mean pitches for each constituent of L/N1, L/N2, R/N2 and R/N3 compounds.



difference in the degree to which pitch drops between these two constituents in R/N2 and L/N2 compounds on the one hand, and R/N3 compounds on the other hand. Whereas pitch drops about 5.5 semitones in L/N2 compounds and 3.5 semitones in R/N2 compounds, respectively, in R/N3 compounds pitch drops only 1 semitone between constituent N2 and N3. Thus, for L/N1 compounds we observe a high pitch on constituent N1 and a clearly lower pitch on constituent N2 and on constituent N3. In contrast, R/N2 compounds have a high pitch on constituent N1 and on constituent N2, with a clearly lower pitch on constituent N3. The pitch pattern of the L/N2 compounds is almost identical to that of R/N2 compounds. Finally, R/N3 compounds have a relatively high pitch on all three constituents due to a constant pitch drop of only 1 semitone between each of the three constituents. Table 6.6 provides the mean pitches and respective standard deviations measured for each constituent in each of the four compound groups.

Interestingly, the pitch patterns observed for the four compound types all appear to be in accordance with the Embedded Prominence Hypothesis (see section 6.1.2 again for details on the predictions). In order to find out whether this visual impression was also statistically supported, I devised a regression model with pitch as dependent variable and BRANCHING and POSITION as independent variable. The inspection of

Table 6.6: Excluded speaker: Mean values and Standard deviations of pitch values for L/N1, L/N2, R/N2 and R/N3 compounds.

Constituent	Mean L/N1	Mean L/N2	Mean R/N2	Mean R/N3
N1	22.34 (SD:2.12)	22.01 (SD:1.92)	21.77 (SD:1.33)	22.77 (SD:2.87)
N2	15.61 (SD:0.85)	22.16 (SD:2.22)	21.29 (SD:2.28)	21.88 (SD:2.41)
N3	16.05 (SD:0.77)	16.53 (SD:1.29)	16.49 (SD:2.93)	20.79 (SD:3.30)

the residuals of this first model revealed one observation with a residual higher a standard deviation of $|3|$. The data point was excluded from the data set and a new model was fitted to the data. A type-III ANOVA of the final model revealed a significant effect for the factor POSITION ($F(2, 251.4) = 27.151$), $p < 0.001$) and a significant interaction of BRANCHING by POSITION ($F(6, 204.0) = 7.3428$, $p < 0.001$). Yet, the main effect for the factor BRANCHING turned out to be non-significant.

The model revealed that the pitch drop between constituent N1 and constituent N2 in L/N1 compounds is significantly larger than in L/N2 compounds ($p < 0.001$) and R/N2 compounds ($p < 0.001$), respectively. In addition to that, the pitch value on constituent N2 in L/N1 compounds turned out to be significantly lower than in L/N2 compounds ($p < 0.001$) and R/N2 compounds ($p < 0.001$).¹⁵ Finally, the pitches assigned to constituent N1 and constituent N3 did not differ significantly between L/N1 compounds on the one hand and R/N2 and L/N2 compounds on the other hand. The result is fully in accordance with the predictions of the Embedded Prominence Hypothesis for L/N1, R/N2 and L/N2 compounds.

In addition to that, the model revealed that pitch drops less steeply between constituent N2 and constituent N3 in R/N3 compounds than in R/N2 compounds ($p < 0.05$). Furthermore, it showed that pitch on constituent N3 is significantly higher in R/N3 compounds than in R/N2 compounds ($p < 0.001$). Moreover, there is no statistical evidence in the data that the two right-branching compound groups differ from each other in pitch assigned to constituent N1 nor in pitch assigned to constituent N2. Hence, this result suggests that R/N3 compounds also behave in accordance with the Embedded Prominence Hypothesis.

¹⁵The summary table of the regression model for the excluded speaker is provided in part B of the Appendix.

Discussion

The previous section presented the result of the pitch analysis for the speaker that was excluded from the data set of this experiment. The analysis of the excluded speaker's data revealed a tendency for all four compound groups to behave as predicted by the Embedded Prominence Hypothesis. The speaker assigned highest pitch to constituent N1 in L/N1 compounds and clearly lower pitches to constituent N2 and constituent N3. In contrast, in R/N2 and L/N2 compounds high pitches were assigned to both constituent N1 and constituent N2. Crucially, the pitch values on constituent N2 were significantly higher in R/N2 and L/N2 compounds than in L/N1 compounds. The result strongly suggests that highest prominence was assigned to constituent N1 in L/N1 compounds and to constituent N2 in R/N2 and L/N2 compounds. With reference to R/N3 compounds, we found that the speaker assigned high pitches to all three constituents, which resulted in a significantly higher pitch on constituent N3 in R/N3 compounds than in R/N2 compounds. The result strongly suggests that highest prominence was generally assigned to constituent N3 in R/N3 compounds by this speaker.

The analysis revealed that the excluded speaker assigned the same prominence pattern to L/N1, L/N2 and R/N2 compounds as the other speakers participating in the experiment, but differed from the remaining 12 speakers in prominence assignment to R/N3 compounds. The fact that the excluded speaker tended to assign highest prominence to constituent N3 in R/N3 compounds might have been one reason why the first model failed to predict the speaker's data appropriately. Another reason for that might have been the speaker's extreme pitch changes between prominent and non-prominent constituents. These are much larger than those obtained for the other speakers. For instance, for L/N1 compounds, we observed that pitch dropped about 6.5 semitones between constituent N1 and constituent N2. In contrast to that for the other 12 speakers, the average pitch fall between these two constituents in L/N1 compounds was found to be only 1.6 semitones high. Similarly, in L/N2 compounds the excluded speaker dropped his pitch about 5.5 semitones between constituent N2 and constituent N3. The respective pitch drop observed for the other speakers amounted to only 1.8 semitones on average.

Finally, we note that the result obtained for the R/N3 compounds for this one particular speaker provides some evidence for the assumption that right-branching compounds with embedded right prominent NN compounds may indeed have highest prominence on constituent N3 as predicted by the EPH. In contrast to the other par-

ticipants, this speaker clearly assigned highest prominence to the complex constituent of right-branching compounds. Thus, whatever factors may have caused the other speakers to vary in prominence assignment to these compounds, they did not affect the excluded speaker.

6.5 Conclusion

The experiment presented in this chapter tested the predictions of the Embedded Prominence Hypothesis. The EPH predicts that if highest prominence falls on the complex constituent of a triconstituent compound, left- and right-branching compounds with embedded left prominent NNs have highest prominence on the left member of the complex constituent. This is the same prominence pattern as predicted by the LCPR. In addition, The EPH also predicts that left- and right-branching compounds with embedded right prominent NNs have highest prominence on the right member of the complex constituent, a prominence pattern not predicted by the LCPR. The analysis of the data showed that left- and right-branching compounds with embedded left prominent NN compounds indeed differ from each other in that highest prominence is assigned to constituent N1 in left-branching compounds and to constituent N2 in right-branching compounds. This result is in accordance with the Embedded Prominence Hypothesis and Liberman and Prince's LCPR.

However, the statistical analysis of the data also revealed that the presence of an embedded right prominent NN compound poses a problem for Liberman and Prince's generalization, at least for left-branching compounds. Left-branching compounds with embedded right prominent NNs were found to have highest prominence on constituent N2, i.e. on the right member of the complex constituent. Although highest prominence is still assigned to one of the two members of the complex constituent in these compounds, the prominence pattern violates the LCPR at the N-level of the left-branching compounds. In contrast, for right-branching compounds with embedded right prominent NNs, the analysis provided no support for the EPH. Instead, the result obtained for these compounds suggests that the speakers varied in prominence assignment to R/N3 compounds, with a tendency to assign highest prominence to constituent N1. However, the analysis of the excluded speaker provides some evidence for the EPH with regard to the group of R/N3 compounds. Because of that, it seems advisable to repeat the present experiment, with newly constructed R/N3 compounds, in order to find out whether the EPH in fact generally fails to account for right-branching compounds with embedded right prominent NNs.

What follows from this experiment is that the Lexical Category Prominence Rule only accounts for the prominence pattern of a specific type of compound rather than functioning as a generalization for the prominence pattern of all NNN compounds. This finding is in line with Giegerich (2009), who also argues that the LCPR is unable to account for all the prominence patterns found among triconstituent compounds. In particular, the LCPR is based on the assumption that NN compounds are generally left prominent as well as on an arbitrary distinction between NN compounds and NN phrases. Yet, if one does not make this distinction between NN compounds and NN phrases, the generalization runs into problems in that it accounts only for the prominence pattern of a restricted number of compounds, i.e. compounds with embedded left prominent NNs. For left-branching compounds with embedded right prominent NNs, the LCPR fails to predict the correct prominence pattern. We therefore conclude that any account that tries to explain prominence assignment in triconstituent compounds needs to incorporate the existence of right prominent NN compounds in English.

7 Production Experiment 2: IC-Prominence Hypothesis (IPH)

This chapter presents a production experiment that tests the IC-Prominence Hypothesis for left-branching NNN compounds. The chapter begins with a description of the stimuli constructed in order to test whether analogical mechanisms and/or semantics trigger left-branching compounds to have highest prominence on constituent N3, contra the LCPR. The description of the stimuli is followed by a section which deals with some preliminary statistical measures. Section 7.2 presents the results of the analysis, which are further discussed in section 7.3. The chapter ends with some concluding thoughts in section 7.4.

7.1 Methodology

7.1.1 Stimuli

Analogy

The question as to whether right prominence at the IC-level of left-branching NNN compounds is triggered by analogy to the head constituent was tested with the two lexical items *avenue* and *pie*. The two lexical items are generally claimed to trigger right prominence in biconstituent NN compounds when functioning as the head constituent of a compound (e.g. Fudge 1984; Ladd 1984; Plag 2003; Bell 2008; Giegerich 2004; Liberman and Sproat 1992). Thus, with each of the two lexical items five left-branching compounds were created in which *avenue* and *pie* functioned as the head constituent. In addition to these five *avenue* compounds and five *pie* compounds, I constructed 10 left-branching compounds which would serve as control group items, i.e. compounds to which the pitch values measured for the *avenue* and *pie* compounds could be compared in the subsequent pitch analysis. Crucially, the control group items had to be clearly left prominent at the IC-level of the compounds. In other words, in

these compounds, highest prominence had to be assigned to one of the two members of the complex modifier, and not to the non-complex head constituent. But how to control for left prominence at the IC-level of the control group items?

With reference to the *avenue* compounds, I decided to compare them to left-branching compounds ending in the lexical item *street*. In the literature on prominence assignment to biconstituent NN compounds (e.g. Fudge 1984; Ladd 1984; Schmerling 1971; Plag 2003), *avenue* and *street* compounds are claimed to clearly contrast with reference to their prominence pattern. In particular, *avenue* compounds are generally claimed to be right prominent (e.g. *Madison Avenue*) whereas *street* compounds favour left prominence (e.g. *Madison Street*). Since the IC-Prominence Hypothesis is based on the assumption that the same factors triggering right prominence in NN compounds also trigger the same prominence behaviour in NNN compounds, I expected that the same prominence contrast observed between biconstituent *avenue* and *street* compounds should also be prevalent in triconstituent compounds. Hence, the five triconstituent compounds with *avenue* as head constituent were compared to five triconstituent compounds with the head constituent *street*.

Furthermore, the five *pie* compounds were compared to left-branching compounds with a deverbal head ending in the suffix *-er*. Empirical studies by Plag et al. (2007, 2008), which tested different hypotheses regarding prominence variation in biconstituent NN compounds, found a strong statistical tendency towards left prominence for this particular subgroup of argument-head compounds (see section 2.1 again for more details). Again, under our assumption that the same mechanisms responsible for prominence assignment in NN compounds may also operate at the IC-level of triconstituent compounds, I expected that the same tendency towards left prominence as observed for NN *-er* compounds is also found among triconstituent left-branching compounds with a deverbal head ending in *-er* (e.g. [*íce hockey*] *player*, [*schóol bus*] *driver*).¹

Importantly, the control group items and the *avenue* and *pie* compounds were constructed as pairs. Thus, for each of the five *avenue* compounds a corresponding *street* compound was created which differed from the *avenue* compound only in its head-constituent (e.g. [*Post office*] *Avenue* vs. [*Post office*] *Street*, [*Fish market*] *Avenue* vs. [*Fish market*] *Street*). The same was true for the five *pie* compounds for each of which a corresponding argument-head compound with the same complex constituent was

¹It should be pointed out again that a statistical tendency towards a certain prominence pattern does not imply that this effect is categorical in nature. Instead, it tells us that the majority of compounds with a deverbal head ending in *-er* tend to be left prominent, yet that there are also some right prominent compounds among this group.

constructed (e.g. [*passion fruit*] *pie* vs. [*passion fruit*] *seller*, [*orange juice*] *pie* vs. [*orange juice*] *supplier*). By keeping the complex constituent constant between the regular test items and the control group items, it was possible to test as to whether the choice of the head constituent causes the compounds to be stressed differently. In particular, it was tested whether *avenue* and *pie* compounds exhibit a higher pitch, i.e. higher prominence, on constituent N3 than the control group items.

Finally, it should be noted that the prominence pattern of the embedded NN compounds was not explicitly controlled when constructing the stimuli as controlling for the prominence pattern of the complex constituent was not considered to be crucial for testing the IC-Prominence Hypothesis. The IC-Prominence Hypothesis predicts that the two lexical items *avenue* and *pie* trigger highest prominence on the head-constituent of a left-branching compound. Thus, *avenue* and *pie* compounds should differ from the control items with respect to the pitch height on constituent N3, irrespective of whether the embedded NN is left or right prominent. In contrast, the control items should be left prominent at the IC-level, i.e. either constituent N1 or constituent N2 should be marked as the most prominent constituent.² For a better illustration of the constructed data the *avenue* and *pie* compounds and their respective control group items are listed in Table 7.1 and Table 7.2.

Table 7.1: List of *avenue* and *street* compounds

avenue compounds	street compounds
[<i>Bird Rock</i>] <i>Avenue</i>	[<i>Bird Rock</i>] <i>Street</i>
[<i>Post office</i>] <i>Avenue</i>	[<i>Post office</i>] <i>Street</i>
[<i>Lemon Grove</i>] <i>Avenue</i>	[<i>Lemon Grove</i>] <i>Street</i>
[<i>Ocean Front</i>] <i>Avenue</i>	[<i>Ocean Front</i>] <i>Street</i>
[<i>Fish market</i>] <i>Avenue</i>	[<i>Fish market</i>] <i>Street</i>

²It is possible that if we embed a right prominent NN compound in a *pie* or *avenue* compound, the prominence pattern of the embedded right prominent NN is changed due to ‘Iambic Reversal’, i.e. a shift of prominence (e.g. Liberman and Prince 1977). However, a shift of prominence caused by Iambic Reversal would only affect the prominence relation between the two members of the complex constituent in that prominence of the second constituent may be shifted to the first constituent of the compound (e.g. Giegerich 2009:9). Thus, despite the possibility of Iambic Reversal constituent N3 should still be the most prominent constituent in *pie* and *avenue* compounds if the IC-Prominence Hypothesis is correct.

Table 7.2: List of *pie* and argument-head compounds

pie compounds	argument-head compounds
<i>[passion fruit] pie</i>	<i>[passion fruit] seller</i>
<i>[candy bar] pie</i>	<i>[candy bar] buyer</i>
<i>[blood orange] pie</i>	<i>[blood orange] supplier</i>
<i>[orange juice] pie</i>	<i>[orange juice] producer</i>
<i>[potato chip] pie</i>	<i>[potato chip] taster</i>

Semantics

The two semantic relations *IC2 is located at IC1* and *IC2 during IC1* were selected in order to test whether the semantics of a compound triggers right prominence at the IC-level of left-branching NNN compounds. According to the literature dealing with prominence assignment in biconstituent compounds, these two semantic relations are assumed to trigger right prominence in NN compounds (e.g. Fudge 1984; Olsen 2000; Bell 2008; Bell and Plag 2012).³ Empirical support for this assumption is provided by Plag et al. (2008) who tested the role of semantics in prominence assignment to biconstituent compounds using a large amount of corpus data. In their investigation of NN compounds taken from the Boston University Radio Speech Corpus, Plag et al. (2008) found a statistical tendency towards right prominence in NN compounds which exhibited one of the above mentioned semantic relations. Based on these findings, it was decided to select the two semantic relations *IC2 is located at IC1* and *IC2 during IC1* to test the same effect for triconstituent compounds.

For each of the two semantic relations selected, I constructed six compounds. Crucially, the compounds had to exhibit the two semantic relations at the IC-level of the compound, i.e. between the complex modifier and the head constituent (e.g. *[holiday season] job*, ‘a job during the holiday season’; *[coffee house] concert*, ‘a concert at a coffee house’). In addition to that, for each of the twelve compounds exhibiting one of the two semantic relations, I created a corresponding control group item, which differed from the semantic item only in its head constituent. Hence, the complex constituent was held constant again in each pair in order to reduce uncontrollable sources of variation as potentially introduced by varying the members of the complex constituent between the control group items and the regular test items. Moreover, it was crucial for the control items to be clearly left prominent at the IC-level, i.e. the most promi-

³Regarding biconstituent compounds the two semantic relations are referred to as *N2 is located at N1* and *N2 during N1*.

Table 7.3: List of *IC2 during IC1* and argument-head compounds

IC2 during IC1	argument-head compounds
<i>[field trip] weather</i>	<i>[field trip] leader</i>
<i>[winter term] breakfast</i>	<i>[winter term] preparation</i>
<i>[Monday morning] meeting</i>	<i>[Monday morning] hater</i>
<i>[flight test] accident</i>	<i>[flight test] manual</i>
<i>[spring break] vacation</i>	<i>[spring break] organization</i>
<i>[holiday season] job</i>	<i>[holiday season] planner</i>

Table 7.4: List of *IC2 is located at IC1* and argument-head compounds

IC2 is located at IC1	argument-head compounds
<i>[cocktail bar] fight</i>	<i>[cocktail bar] designer</i>
<i>[coffee house] concert</i>	<i>[coffee house] lover</i>
<i>[gas station] robbery</i>	<i>[gas station] operator</i>
<i>[apartment building] party</i>	<i>[apartment building] manager</i>
<i>[hockey stadium] event</i>	<i>[hockey stadium] cleaner</i>
<i>[movie theater] fire</i>	<i>[movie theater] visitor</i>

ment constituent of the whole compound should be constituent N1 or constituent N2. Therefore, the head constituents of the control group items were chosen in such a way that when combining the potential head with the complex constituent the resulting compound would exhibit an argument-head relation at its IC-level (e.g. *hockey stadium + cleaner = hockey stadium cleaner*).

Finally, as pointed out for the *avenue* and *pie* compounds, the prominence pattern of the complex constituent was not explicitly controlled for either left- or right prominence as there was no reason to assume that the prominence pattern of the embedded NN compound would affect prominence assignment at the IC-level of the compounds. For a better illustration of the constructed compounds, the *IC2 is located at IC1* compounds and their respective control items are given in Table 7.3; the *IC2 during IC1* compounds and the corresponding control items are given in Table 7.4.

According to the IC-Prominence Hypothesis, the compounds exhibiting one of the two semantic relations at the IC-level should be assigned highest prominence to constituent N3. Opposed to that, the control group items should have highest prominence on constituent N1 or N2. Thus when comparing the pitch values of the two groups, we would expect speakers to assign higher pitches to constituent N3 in the test items than in the control group items. With reference to the pitch values assigned to constituent N1 and constituent N2, we would not expect the *IC2 is located at IC1* and *IC2*

during IC1 compounds to differ from their control group items. The reason for this assumption is that the regular test items and the control items share the same complex constituents.

7.1.2 Statistical procedure

As described in subsection 7.1.1 I constructed 4 different subsets in order to test the IC-Prominence Hypothesis for the factors semantics and analogy. Because of that, I decided to devise four separate mixed-effects models: one model for each of the two lexical items and one model for each of the two semantic relations under investigation. The dependent variable in each of the four models was pitch in semitones (pitchST). In addition to that, I fitted two random effects to each model. The first random effect was that of SPEAKER. By modelling the speakers participating in the experiment as a random effect I tried to account for variance introduced by speaker related differences such as differences in the speakers' relative pitch. As a second random effect, I included the complex constituent of the compounds. In that the four models of the present experiment differ from the previous models presented in this thesis in which the entire item was treated as a random factor. However, there are two reasons for including only the complex constituent as a random effect in each of the four models instead of the entire item as for instance done in experiment 1.

First, as described in the previous section the triconstituent compounds in the present experiment were constructed as pairs. The two subsets for the factor analogy contain 5 such pairs, the subsets for the semantic relations contain 6 pairs (see Tables 7.1; 7.2; 7.3; 7.4). Every two compounds in each subset have the same complex constituent, but differ in their head constituent from each other. The head constituent is the crucial element in these compounds, as - depending on the choice of that constituent - the compound is expected to exhibit highest prominence either on constituent N3 or on one of the two other constituents of the compound, i.e. N1 or N2. Thus, the variance introduced by the choice of a different head constituent is part of the tested hypothesis and thus must be captured in the fixed effects structure of the model rather than become part of the random effects structure (see below for more details on the fixed effects of the model). In contrast, the complex constituent should be included as a random effect in each model in order to account for the variance introduced by the different complex constituents within each subset (e.g. *Post office* vs. *Fish market* vs. *Bird rock*). The random effect capturing the information about the complex

constituents of the compounds is referred to as COMPLEX.⁴

In addition to the two random factors, I added the two factors POSITION and GENDER as fixed factors to each model. The factor GENDER had the two factor levels *male* and *female* and was added to the model in order to account for differences in the pitch range between male and female speakers. The second factor POSITION consisted of the three factor levels *lpitch*, *mpitch* and *rpitch*. Analogous to the corpus study and experiment 1, the three levels of the factor POSITION refer to the three constituents of the compounds for which pitch was measured. In addition to GENDER and POSITION, the four models also contained a fourth factor, which functioned as a grouping factor for the regular test items and the control items. This fourth factor consisted of two factor levels, which represented the regular test items and the control items. In the *avenue* and *pie* model this factor was labelled ANALOGY, with the two factor levels *avenue* and *street* on the one hand and *pie* and *arg-head* on the other hand. In contrast, in the two semantic models, this grouping factor was labelled SEMANTICS and the compounds exhibiting the semantic relations *IC2 during IC1* and *IC2 is located at IC1* were represented by the factor levels *duration* and *location*, respectively; the control group items were labelled *control* in each of the two models fitted to the semantic subsets.

Finally, a three-way interaction of ANALOGY by POSITION by GENDER and SEMANTICS by POSITION by GENDER, respectively, was included in each of the four models in order to test for potential relations between the three factors. The interaction parameter for the three-way interaction was formulated in a way that all subordinated two-way interactions were tested for significance as well. This was important since the IC-Prominence Hypothesis predicts a significant interaction of ANALOGY by POSITION and SEMANTICS by POSITION, respectively. In particular, the presence of a significant interaction of ANALOGY/SEMANTICS by POSITION would imply that the regular test items and the control items differ in their pitch pattern from each other. In contrast, the absence of this interaction would suggest that speakers did not assign different pitch patterns to the two compound groups and thus did not assign different prominence patterns to the control and regular test items.

⁴That the inclusion of SPEAKER and COMPLEX as random effects was justified in the four models was verified by means of likelihood ratio tests. See Appendix part B.

7.2 Results

This section presents the results of the statistical analysis of the data investigated in the present experiment. The section begins with the results obtained for the two lexical items *avenue* and *pie* and continues with the results for the two semantic relations *IC2 during IC1* and *IC2 is located at IC1*.

7.2.1 Analogy: *avenue* vs. *street*

The subset of the *avenue* and *street* compounds contained a total of 465 pitch measurements.⁵ A mixed-effects model with the model specifications described in the previous section was fitted to the data. During subsequent model criticism all data points with residuals exceeding a standard deviation of |2.5| were excluded (N = 13, 2.8%) from the data and a new model was refitted to the trimmed data set. By means of model comparison it was investigated whether higher order interactions could be removed from the model without reducing the model's explanatory power. This procedure resulted in the exclusion of the three-way interaction of ANALOGY by POSITION by GENDER. Furthermore, all subordinated two-way interactions involving GENDER could be dropped from the model, with GENDER only remaining as a main effect in the model. This result tells us that male and female speakers did not differ in prominence assignment to any of the constructed compounds, but that the two genders again only differed with respect to the overall pitch height. Crucially, the interaction of ANALOGY by POSITION turned out to be significant. This significant interaction of ANALOGY by POSITION suggests that either between constituent N1 and constituent N2 or between constituent N2 and constituent N3, the speakers dropped their pitch to a significantly different degree in *avenue* and *street* compounds. In order to know at which position the two groups differ from each other, however, one needs to look at the table of coefficients of the model.

The table of coefficients of the final *avenue* and *street* model is given in Tables 7.5-7.7. In each table, the same model is provided with a different baseline, so that all pairwise comparisons relevant to answering the IC-Prominence Hypothesis are directly displayed in one of the three tables. The baseline of the model in Table 7.5 is that of *lpitch, avenue and female*, i.e. the estimated mean pitch of constituent N1 in *avenue*

⁵10 data points from the female portion of the *avenue* vs. *street* subset were excluded prior to fitting the mixed-effects model to the data. The values were clearly lower than the remaining values and a closer inspection of the data points revealed that these were instances of creaky voice, which had not been excluded before.

Table 7.5: Table of Coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: ANALOGY = avenue, POSITION = lpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	23.8721	0.7853	30.398	< 0.001
mpitch	-1.1028	0.1657	-6.656	< 0.001
rpitch	-2.1880	0.1670	-13.106	< 0.001
street	0.0826	0.1607	0.514	0.608
male	-8.6970	1.1474	-7.580	< 0.001
mpitch:street	-0.6433	0.2303	-2.793	0.005
rpitch:street	-0.6042	0.2346	-2.575	0.010

Table 7.6: Table of Coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: ANALOGY = avenue, POSITION = mpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	22.7693	0.7859	28.974	< 0.001
lpitch	1.1028	0.1657	6.656	< 0.001
rpitch	-1.0853	0.1691	-6.416	< 0.001
street	-0.5608	0.1654	-3.389	< 0.001
male	-8.6970	1.1473	-7.580	< 0.001
lpitch:street	0.6433	0.2303	2.793	0.005
rpitch:street	0.0391	0.2378	0.165	0.869

Table 7.7: Table of Coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: ANALOGY = avenue, POSITION = rpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	21.6840	0.7860	27.588	< 0.001
lpitch	2.1880	0.1670	13.106	< 0.001
mpitch	1.0853	0.1691	6.416	< 0.001
street	-0.5216	0.1714	-3.043	0.003
male	-8.6970	1.1474	-7.580	< 0.001
lpitch:street	0.6042	0.2346	2.575	0.010
mpitch:street	-0.0391	0.2378	-0.165	0.869

compounds for female speakers. In contrast, in Table 7.6 the baseline represents the mean pitch estimated for constituent N2 in *avenue* compounds for female speakers. Finally, the selected baseline in Table 7.7 is the mean pitch of constituent N3 in *avenue* compounds.

For a visual illustration of the result, the interaction plot of the final model is given in Figure 7.1. Pitch in semitones is given on the vertical axis, the three pitch positions for which pitch was measured in the compounds are displayed on the horizontal axis. The solid line represents the group of *avenue* compounds, the dashed line that of *street* compounds. The lines connect the average mean pitches (circles) estimated for each position in the *avenue* and *street* compounds for the female portion of the data set. The corresponding pitch values for male speakers are generally 8.697 semitones lower than the estimated pitch values for the female speakers (see Tables 7.5-7.7).

Figure 7.1: Interaction plot displaying the average mean pitches estimated by the mixed-effects model for each constituent in the *avenue* and *street* compounds; female speakers.

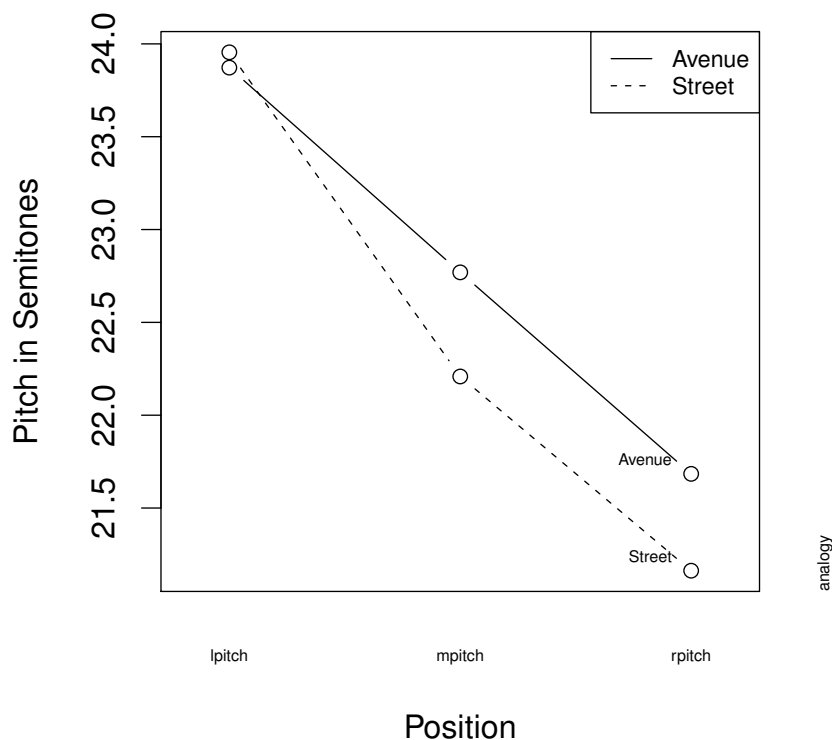


Figure 7.1 shows that in both *avenue* and *street* compounds pitch generally decreases from constituent N1 to constituent N3. This general downtrend of pitch may be re-

garded as the result of a general pitch declination effect (e.g. Gussenhoven 2004). Furthermore, the plot shows that *avenue* and *street* compounds differ in the degree to which pitch drops over the course of the compound utterances. At first, we observe that pitch assigned to constituent N1 is the same in *avenue* and *street* compounds (see Table 7.5, *street*, $p = 0.608$). However, due to a significantly larger pitch drop between constituent N1 (lpitch) and constituent N2 (mpitch) in *street* compounds (see Table 7.5, mpitch:Street, $p = 0.005$), the two compound groups differ in pitch assigned to constituent N2. In particular, in *avenue* compounds pitch on constituent N2 is on average about 0.56 ST higher than in *street* compounds (see Table 7.7, *street*, $p < 0.001$). The pitch difference is indicated by the relatively large gap between the circles at position 'mpitch' in the plot. More important with reference to the IC-Prominence Hypothesis, however, is the fact that *avenue* compounds also tend to have a higher pitch (+ 0.52 ST) on constituent N3 than *street* compounds (see Table 7.6, *street*, $p = 0.003$). The plot shows that pitch drops to the same degree between constituent N2 and constituent N3 in *avenue* and in *street* compounds due to which the higher pitch on constituent N2 in *avenue* compounds seems to be maintained by the speakers on constituent N3. This higher pitch on constituent N3 in *avenue* compounds is in accordance with the IC-Prominence Hypothesis.

Thus, according to the analysis speakers assigned different pitch patterns to *avenue* and *street* compounds. In *street* compounds speakers assigned highest pitch to constituent N1 and clearly lower pitches to constituent N2 and constituent N3. In contrast, *avenue* compounds exhibit a relatively high pitch on all three constituents. Crucially, *avenue* compounds tend to have a higher pitch on constituent N3 than *street* compounds which is in accordance with the IC-Prominence Hypothesis. The result that *avenue* and *street* compounds also differ in pitch on constituent N2 is rather unexpected given that the *avenue* and *street* compounds shared the complex constituent with each other. The result is discussed in more detail in section 7.3.

7.2.2 Analogy: *pie* vs. argument-head compounds

The *pie* subset contained a total of 461 pitch measurements. 8 pitch measurements were excluded after a first separate inspection of the pitch distribution of the female and male subsets; these values clearly stood out from the rest in that they were extremely low in comparisons to the remaining values. As it turned out these values were again either due to use of non-modal phonation or bad recording quality. A mixed-effects model was fitted to the remaining 453 observations. During subsequent

model criticism data points with residuals higher a standard deviation than $|2.5|$ were excluded from the data ($N = 14, 3.1\%$) and the model was refitted to the trimmed data set. By means of model comparison, it was checked whether the inclusion of all parameters mentioned in section 7.1.2 was fully justified. Following the procedure described, for example, in Crawley (2005b) and Baayen (2008), all non-significant parameters were excluded from the model in a step-wise fashion.

The table of coefficients of the final *pie* vs. argument-head model is given in Tables 7.8-7.10. As for the *avenue* vs. *street* model, each of the three tables has a different position as its baseline. That way all relevant pairwise comparisons between *pie* and argument-head compounds are directly displayed in one of the three tables. Table 7.8 has the baseline *lpitch, pie* and *female*. In Table 7.9, the baseline was altered to *mpitch, pie* and *female*. The selected baseline in Table 7.10 is that of the *rpitch, pie* and *female*.

The three tables of coefficients show that only the interaction of ANALOGY by POSITION survived in the model, with the three-way interaction of ANALOGY by GENDER by POSITION and the two-way interactions of GENDER by POSITION and ANALOGY by GENDER being removed from the model. The factor GENDER survived as a main effect in the model. This tells us that male and female speakers did not assign different prominence patterns to any of the constructed compounds, but only differed with reference to the general pitch height. In particular, according to the model the pitch used by male speakers is on average about 8 semitones lower than that of the female speakers. The significant interaction of ANALOGY by POSITION tells us that *pie* compounds and argument-head compounds differ significantly from each other in at least one of the pitch relations between the three compound constituents.

For a better illustration of the result, the interaction plot of the model is given in Figure 7.2. Analogous to the previous interaction plots, pitch in semitones is given on the vertical axis, the three pitch positions on the horizontal axis. The *pie* compounds are represented by the solid line and the argument-head compounds by the dashed line.

Table 7.8: Table of Coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: ANALOGY = pie, POSITION = lpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	23.1923	0.8544	27.146	< 0.001
mpitch	-0.6666	0.2038	-3.27	< 0.001
rpitch	-1.9899	0.2031	-9.80	< 0.001
arg-head	0.2307	0.1979	1.17	0.245
male	-8.1744	1.29	-6.318	< 0.001
mpitch:arg-head	-0.7791	0.2851	-2.733	0.007
rpitch:arg-head	-0.4175	0.2841	-1.470	0.142

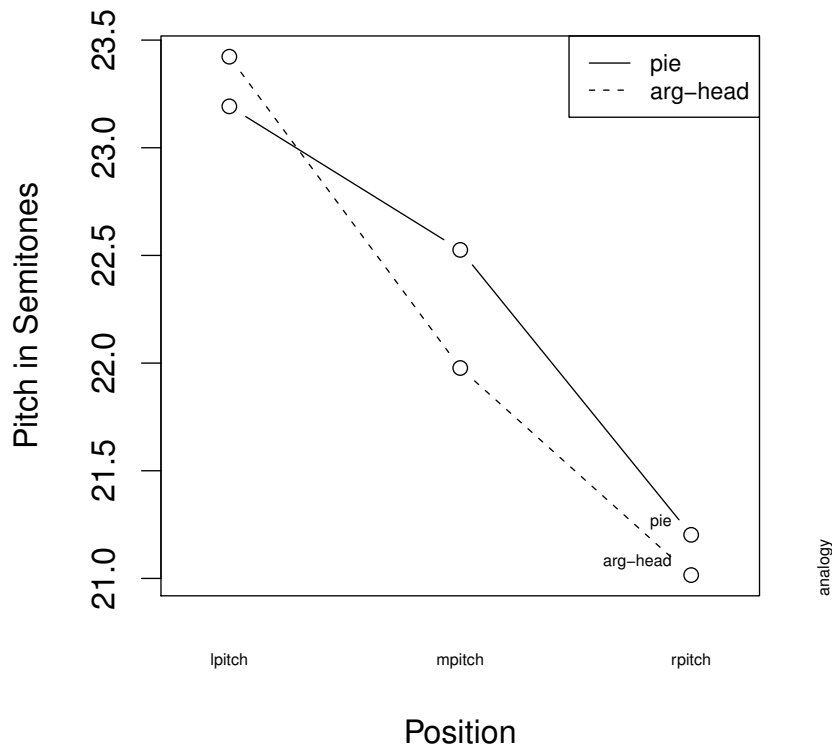
Table 7.9: Table of Coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: ANALOGY = pie, POSITION = mpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	22.5256	0.8556	26.328	< 0.001
lpitch	0.6666	0.2038	3.271	< 0.001
rpitch	-1.3233	0.2077	-6.371	< 0.001
arg-head	-0.5484	0.2059	-2.663	0.008
male	-8.1744	1.2938	-6.318	< 0.001
lpitch:arg-head	0.7791	0.2851	2.733	0.007
rpitch:arg-head	0.3616	0.2898	1.248	0.213

Table 7.10: Table of Coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: ANALOGY = pie, POSITION = rpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	21.2024	0.8556	24.780	< 0.001
lpitch	1.9899	0.2031	9.799	< 0.001
mpitch	1.3233	0.2077	6.371	< 0.001
arg-head	-0.1869	0.2047	-0.913	0.362
male	-8.1744	1.2938	-6.318	< 0.001
lpitch:arg-head	0.4175	0.2841	1.470	0.142
mpitch:arg-head	-0.3616	0.2898	-1.248	0.213

Figure 7.2: Interaction plot displaying the average mean pitches estimated by the mixed-effects model for each constituent in the *pie* and argument-head compounds; female speakers.



A look at Figure 7.2 shows that in both *pie* and argument-head compounds highest pitch is assigned to constituent N1. Furthermore, according to our model argument-head and *pie* compounds do not differ in pitch on constituent N1 (see Table 7.8, arg-head = 0.2307, $p = 0.245$). However, we observe that *pie* compounds tend to have a higher pitch on constituent N2 than argument-head compounds. This difference is due to a smaller pitch drop between constituent N1 and constituent N2 in *pie* compounds. According to Table 7.8 this difference in the pitch relation between constituent N1 and constituent N2 of *pie* and argument-head compounds is also statistically highly significant (mpitch:arg-head = -0.7791 , $p < 0.007$). This is also the case for the difference between the mean pitches estimated for constituent N2 in *pie* and argument-head compounds (see Table 7.9, arg-head = -0.5484 , $p = 0.008$). Thus, speakers assigned a statistically significantly higher pitch to constituent N2 in *pie* compounds than in argument-head compounds. However, contrary to the IC-Prominence Hypothesis, we also observe a slightly steeper pitch drop between constituent N2 and constituent N3 in *pie* compounds than in argument-head compounds. Although Table

7.10 shows that this difference in the pitch change from constituent N2 to constituent N3 between the two groups does not reach significance (mpitch:arg-head = -0.3616 , $p = 0.231$), the two mean pitches on constituent N3 differ only slightly between the two groups. In fact, Table 7.10 reveals that the p-value of the coefficient capturing this difference between the mean pitch on constituent N3 is non-significant (arg-head = 0.3616 , $p = 0.362$). Thus, a significant difference between pitch on constituent N3 in *pie* and argument-head compounds is not traceable, a result which is contra the IC-Prominence Hypothesis.

The analysis of the present subset revealed that *pie* compounds exhibit a higher pitch on constituent N2 than argument-head compounds. Furthermore, the two groups neither differed in pitch assigned to constituent N1 nor in that assigned to constituent N3. This result is against the IC-Prominence Hypothesis, according to which we would have expected a clearly higher pitch on constituent N3 in *pie* compounds than in argument-head compounds. However, although the IC-Prominence Hypothesis is not supported by the data, we observed that *pie* and argument-head compounds behave differently. The pitch pattern obtained for the *pie* compounds resembles the pitch pattern observed for left-branching compounds with highest prominence on constituent N2 in experiment 1, whereas that of the argument-head compounds is similar to the ones observed for left-branching compounds with highest prominence on constituent N1 in the corpus study and in experiment 1. This may suggest that speakers tended to assign highest prominence to constituent N2 in the *pie* compounds, but to constituent N1 in the argument-head compounds. However, given the way the compounds were constructed this result is unexpected; *pie* compounds and argument-head compounds shared the same complex constituents and thus in the absence of an analogical effect, we would have expected the two groups to exhibit the same pitch pattern, i.e. prominence pattern. The result is discussed in more detail in section 7.3. Yet before turning to the discussion of the result, we have a look at the results for the two semantic relations.

7.2.3 Semantics: *IC2 during IC1*

The data set of the *IC2 during IC1* compounds contained 570 pitch measurements.⁶ A mixed-effects model with the model specifications described in section 7.1.2 was fitted to the data set. During subsequent model criticism it turned out that data by

⁶10 observations were excluded a priori to the statistical modelling after a first inspection of the female and male subset. Again by excluding these data points I wanted to avoid extreme values to become overly influential in the final analysis.

one particular speaker caused a problem for the model, even after the exclusion of residuals exceeding a standard deviation of $|2.5|$. It was decided to exclude the data points of that particular speaker ($N=33$) from the *IC2 during IC1* data subset in order to avoid misspecifications of the model. A new model was fitted to the trimmed data set. The process of model criticism was repeated for the trimmed data set and a new model was refitted to the remaining data.⁷

During subsequent model comparison the three way-interaction of GENDER by SEMANTICS by POSITION and the two-way interaction of GENDER by SEMANTICS turned out to be non-significant and was thus excluded from the model. Yet, what is more important with reference to the IC-Prominence Hypothesis is the fact that during this process it also turned out that the interaction of SEMANTICS by POSITION was non-significant. The absence of a significant interaction of SEMANTICS by POSITION suggests that *IC2 during IC1* compounds and the corresponding control items do not differ significantly from each other in how pitch changes between the three compound constituents. In other words, the data does not provide any evidence that speakers assigned different prominence patterns to *IC2 during IC1* compounds and the control items.

Table 7.11 gives the table of coefficients of the final *IC2 is during IC1* vs. control items model. I decided to keep the interaction parameter of SEMANTICS by POSITION in the final model despite its non-significance. The reason for documenting a model with a non-significant interaction is that this interaction was part of the hypothesis that was tested in the present experiment. In addition to that, the model contains the interaction of GENDER by POSITION, which was found to be significant during model simplification.⁸ Finally, the factor SEMANTICS remained in the model as a main effect. The baseline of the table of coefficients given in Table 7.11 is that of *mpitch, duration and female*, i.e. the mean pitch estimated for constituent N2 for female speakers in *IC2 is during IC1* compounds. By selecting this particular baseline, the pitch differences between constituent N1 and constituent N2 on the one hand, and constituent N2 and constituent N3 on the other hand in *IC2 during IC1* compounds can be directly read off the table. Furthermore, despite the absence of a significant interaction of POSITION by SEMANTICS, the interaction plot of the model is given in Figure 7.3 for a better illustration of this result.

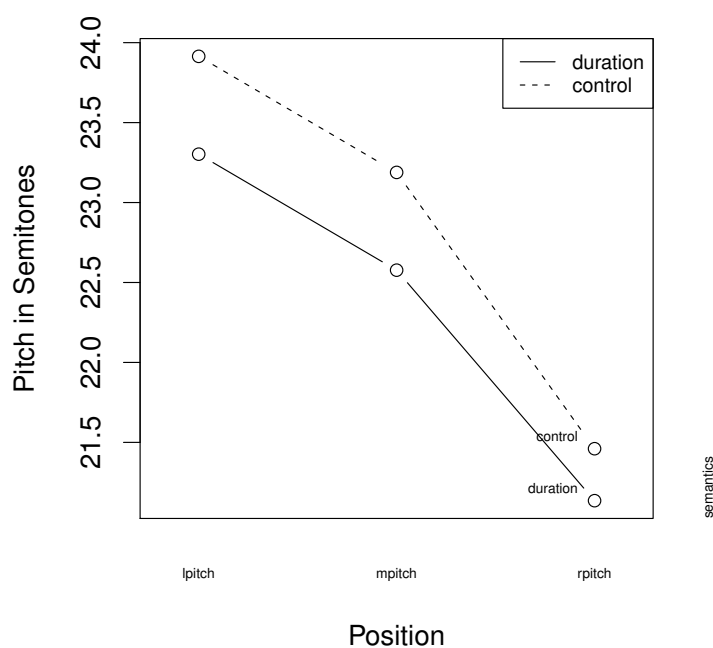
⁷Another 11 data points were excluded from the data set due to residuals exceeding a standard deviation of $|2.5|$, which reduced the number of observations to 526.

⁸The interaction of GENDER by POSITION is irrelevant with reference to the research question at hand and thus not discussed in more detail. Yet, it was kept in the model in order to account for the variance explained by it.

Table 7.11: Table of Coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: SEMANTICS = duration, POSITION = mpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	22.5770	0.7864	28.709	< 0.001
lpitch	0.7255	0.6563	4.380	< 0.001
rpitch	-1.4417	0.1709	-8.438	< 0.001
control	0.6115	0.1406	4.349	< 0.001
male	-8.8080	1.1693	-7.533	< 0.001
lpitch:control	0.0000	0.1976	0.002	0.998
rpitch:control	-0.2864	0.2034	-1.408	0.160
lpitch:male	-0.1019	0.1991	-0.512	0.609
rpitch:male	0.7261	0.2054	3.535	< 0.001

Figure 7.3: Interaction plot displaying the average mean pitches estimated by the mixed-effects model for each constituent in the IC2 *during* IC1 compounds and *control* items; female speakers.



The *IC2 during IC1* compounds are represented by the solid line whereas the dashed line connects the means of the control items.

The plot shows that in both the *IC2 during IC1* and the control items pitch drops less steeply between constituent N1 and constituent N2 than between constituent N2 and constituent N3. Furthermore, we observe a slightly larger pitch drop between constituent N2 and constituent N3 in the control items than in the test items, yet as noted in the previous paragraph, the difference is statistically non-significant according to the model ($\text{rpitch:control} = -0.2864$, $p = 0.160$). Hence, the two groups behave statistically in the same way, except that the control items generally yield higher pitches on each of the three constituents than the *IC2 during IC1* compounds (control = 0.6115, $p < 0.001$).

To sum up, the result of the present analysis is not in accordance with the IC-Prominence Hypothesis. The *IC2 during IC1* compounds do not exhibit a higher pitch on constituent N3 than the control items. Instead, the two groups exhibit the same pitch pattern, i.e. speakers assigned relatively high pitches to constituent N1 and constituent N2 and clearly lower pitches to constituent N3. The pitch pattern of the two compound groups is similar to the pitch pattern observed for compounds with highest prominence on constituent N2 in the corpus study and experiment 1. This may indicate that highest prominence was also assigned to constituent N2 in the two compound groups analyzed in the present analysis, although I do not want to rely too heavily on this interpretation. Finally, speakers generally tended to use higher pitches on the three constituents in the control items than in the *IC2 during IC1* compounds. Given the fact that the control items and *IC2 during IC1* compounds shared the same complex constituents, this observation seems unexpected and will be discussed in more detail in section 7.3.

7.2.4 Semantics: *IC2 is located at IC1*

The subset of the *IC2 is located at IC1* compounds contained 568 pitch measurements.⁹ After a first model was fitted to the data, the residuals were investigated and those exceeding a standard deviation of $|2.5|$ were excluded (N=16, 2.8%). A new model was refitted to the trimmed data set (N = 552). During subsequent model comparisons it turned out that only the interaction of GENDER by POSITION remained significant. This result suggests that *IC2 is located at IC1* compounds do not differ significantly from the control items in the degree to which pitch changes between each of the three constituents. In addition to that, it was also possible to remove the factor SEMANTICS from the model without reducing the model's explanatory power. It follows from this that the control items and the *IC2 is located at IC1* compounds also do not differ significantly from each other in any of the three mean pitches estimated by the model. Thus, according to the analysis the *IC2 is located at IC1* compounds and the control items have the same pitch pattern. This result is contra the IC-Prominence Hypothesis.

The final model of the *IC2 is located at IC1* vs. control items is documented in Table 7.12. Despite their non-significance, the interaction of SEMANTICS by POSITION and the factor SEMANTICS were kept in the final model since both parameters were part of the tested hypothesis. The baseline represents the estimated mean pitch for constituent N2 (`mpitch`) of female speakers for *IC2 is located at IC1* compounds. The respective interaction plot of the model is given in Figure 7.4 for a better illustration of the result.

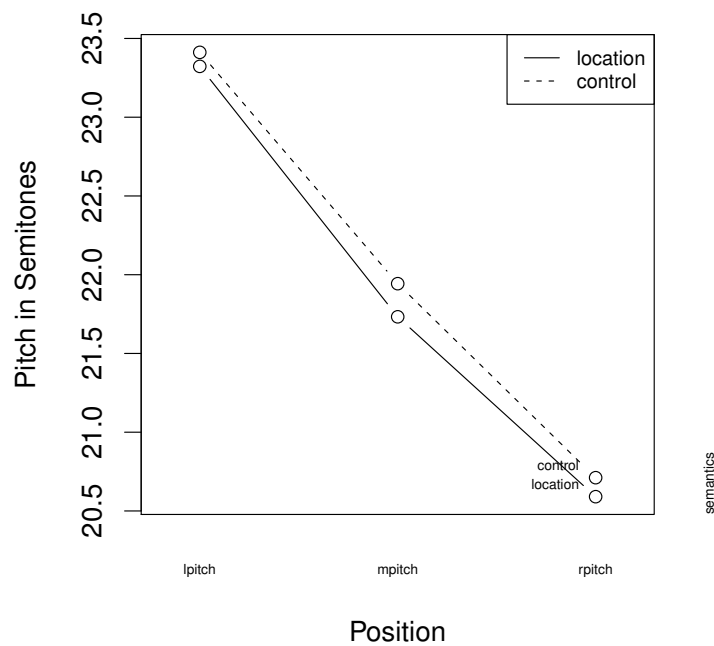
The plot shows that pitch generally decreases over the utterance of the *IC2 is located at IC1* compounds and the control group items, yet with a slightly larger pitch drop between constituent N1 and constituent N2 (1.59 ST) than between constituent N2 and constituent N3 (1.14 ST). Furthermore, the plot shows that the lines connecting the mean pitches of the *IC2 is located at IC1* compounds and the control group items are parallel, which illustrates the absence of a significant interaction of SEMANTICS by POSITION. Furthermore, the fact that the two lines nearly overlap illustrates the absence of a main effect for the factor SEMANTICS. Finally, we note that the pitch pattern of the *IC2 is located at IC1* compounds and their control items is quite similar to that observed for L/N1 compounds analysed in experiment 1. This suggests that highest prominence is assigned to constituent N1 in the two compound groups.

⁹An inspection of the female and male subsets revealed 10 data points (4 = male, 6 = female) which clearly stood out from the rest in that they were either extremely low or high. Hence, I decided to remove those values a priori to the statistical modelling of the data in order to prevent these extreme values to become overly influential in the subsequent analysis. This reduced the number of observations to 568.

Table 7.12: Table of coefficients displaying the fixed effects coefficients of the final mixed-effects model. Baseline: SEMANTICS = location, POSITION = mpitch, GENDER = female

	Estimate	Std. Error	t-value	p-value
(Intercept)	21.7328	0.8678	25.043	< 0.001
lpitch	1.5901	0.1944	8.179	< 0.001
rpitch	-1.1422	0.2017	-5.662	< 0.001
control	0.2103	0.1668	1.261	0.589
male	-8.3139	1.3368	-6.219	< 0.001
lpitch:control	-0.1213	0.2341	-0.518	0.605
rpitch:control	-0.0897	0.2387	-0.376	0.894
lpitch:male	-0.4675	0.2372	-1.971	0.049
rpitch:male	0.2728	0.2406	1.134	0.002

Figure 7.4: Interaction plot displaying the average mean pitches estimated by the mixed-effects model for each constituent in the *IC2 is located at IC1* compounds and *control* items; female speakers.



7.2.5 Individual item variation

The analysis of the *IC2 is during IC1* compounds and *IC2 is located at IC1* compounds revealed no effect in favour of the IC-Prominence Hypothesis. Thus, neither the semantic relation of *IC2 is during IC1* nor that of *IC2 is located at IC1* generally triggered the respective compounds to exhibit highest prominence on constituent N3. However, it turned out that the control items of the *IC2 is during IC1* compounds had consistently higher pitches than the regular test items¹⁰, a result which is rather surprising given that control and regular test items only differed in their head constituent from each other. In addition to that, a closer inspection of the diagnostic plots of the two semantic models revealed that even after the exclusion of observations with high residuals, the final model failed to account for a number of values at the lower and higher end of the distribution (see Figures 7.5 and 7.6). It may be possible that these data points belong to compounds that actually behave according to the IC-Prominence Hypothesis, but which are nevertheless treated as outliers as the majority of the data behaves differently. Hence, in order to gain a better understanding of the semantic data, I decided to explore the individual item pairs in more detail.

Figure 7.5: Quantile-quantile plot for the residuals of the *IC2 during IC1* model.

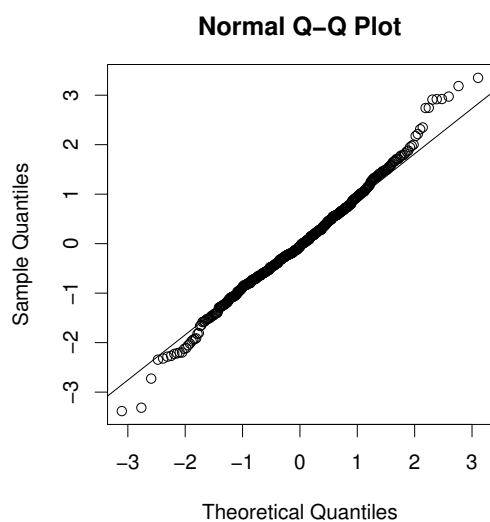
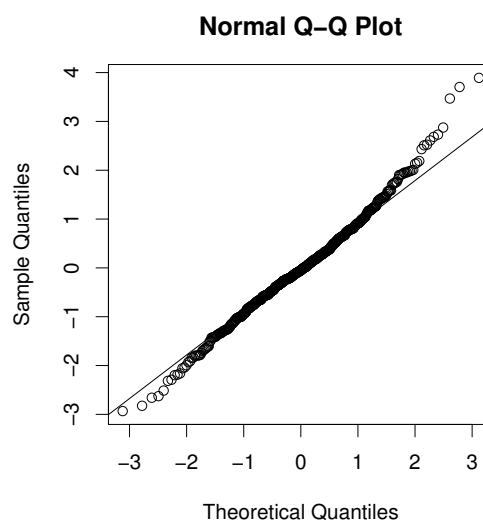


Figure 7.6: Quantile-quantile plot for the residuals of the *IC2 is located at IC1* model.



For each of the 6 pairs of the *IC2 is during IC1* data set as well as for each of the

¹⁰When I speak of 'test items' in this subsection I refer to the *IC2 during IC1* and *IC2 located at IC1* compounds, respectively. The control compounds of each subset are referred to as 'control (group) items'.

6 pairs of the *IC2 is located at IC1* data set, I devised a separate mixed-effects model. In each model I included the factor *SPEAKER* as random effect and the factors *POSITION*, *SEMANTICS* and *GENDER* as fixed effects. In addition, I added a parameter for a two-way interaction of *SEMANTICS* by *POSITION* to each model as the presence of this interaction is predicted by the IC-Prominence Hypothesis. In contrast to the two previous models in this chapter, the complex constituent was not fitted as a random effect to the 12 different models because the compounds compared to each other in these models did not differ regarding the complex constituent (e.g. *hockey stadium event* vs. *hockey stadium cleaner*). Furthermore, the factor *GENDER* was only fitted as a main effect to each of the twelve models in order to account for the general difference in the pitch range between male and female speakers.

The decision not to include this factor in any additional interaction parameter and thus to control for potential prominence differences between male and female speakers was primarily due to methodological reasons. The number of observations in each subset was quite small, with an average of 86 observations per model. Given this small number of observations in each model, one runs the danger of overfitting the model when adding too many predictors to it. In addition to that, the previous models in which potential gender differences were controlled for revealed no significant pitch differences, i.e. prominence differences, between male and female speakers. The only interaction found to be significant was that of *POSITION* by *GENDER*, which is irrelevant with respect to the tested hypothesis.

In each model the residuals were checked for potential outliers and data points with residuals exceeding a standard deviation of $|2.5|$ were excluded.¹¹ Furthermore, by means of model comparison it was checked whether the fixed effects were statistically significant. I begin the exploration of the semantic data with the groups of *IC2 is during IC1* compounds.

IC2 during IC1

Figures 7.7 - 7.12 show 6 interaction plots, i.e. one plot for each model. Thus, each interaction plot displays one of the 6 constructed test items together with their corresponding control items (for an overview of the constructed items see again Table 7.3 in section 7.1.1). In each plot pitch in semitones is given on the vertical axis; the three pitch positions are displayed on the horizontal axis. The average pitch means for the test items are connected by the solid line labelled 'duration', whereas the respective

¹¹The average number of excluded data points in each model was two.

means of the control items are connected by the dashed line labelled ‘control’. The mean values displayed in the plot are those estimated for an average female speaker. The corresponding values for the male speakers are about 8 semitones lower in each model.

A first glance at the six interaction plots reveals that the pitches of the control items are generally higher than those of the test items, which is in line with the general result observed for the *IC2 during IC1* subset. The only exception with respect to this observation is the ‘rpitch’ of the control items of the *spring break* subset, as shown in Figure 7.12. For this item pair we observe that the test items tend to have a higher pitch on constituent N3 than the respective control items. Furthermore, we observe that for the four subsets shown in Figures 7.7, 7.8, 7.9 and 7.10 the pitch change between constituent N1 and constituent N2 is generally smaller than the pitch change between constituent N2 and constituent N3. This holds for both the control and test items. In contrast, Figures 7.10 and 7.11 show that the test items of these two subsets as well as the control items of the *monday morning* subset have a larger pitch change between constituent N1 and constituent N2 than between constituent N2 and constituent N3. Regarding the control items of the *spring break* subset, we note that pitch drops to the same degree between all three constituents.

Interestingly, the four plots displaying the results for the *holiday season*, *winter term*, *flight test* and *spring break* compounds all show clear pitch differences between the test items and their respective control items with respect to the degree to which pitch changes between at least two of the three constituents. In particular, while for the *holiday season* and *winter term* subsets shown in Figures 7.7 and 7.8, we observe that the control items tend to have a higher pitch on constituent N1 than the test items ($p < 0.001$), the control items of the *flight test* subset tend to have a higher pitch on constituent N2 than the test items of the same subset ($p < 0.001$). In addition to that, the *holiday season* plot also shows a difference between the control and test items regarding the mean pitches estimated for constituent N3 ($p < 0.001$).¹²

Given these pitch differences between the control and test items of the 6 subsets, we may raise the question whether these differences also indicate different prominence patterns. The average pitch pattern of the test items of the *holiday season* and *winter term* subsets strongly suggest that highest prominence is assigned to constituent N2 in these compounds. This is indicated by the equally high pitches on constituent N1

¹²The mixed-effects model of each individual item pair for which I determined statistically significant differences are documented in part B of the Appendix. The mixed-effects models of item pairs for which I determined no significant interaction are not documented in the Appendix.

Figure 7.7: [*holiday season*] + N3

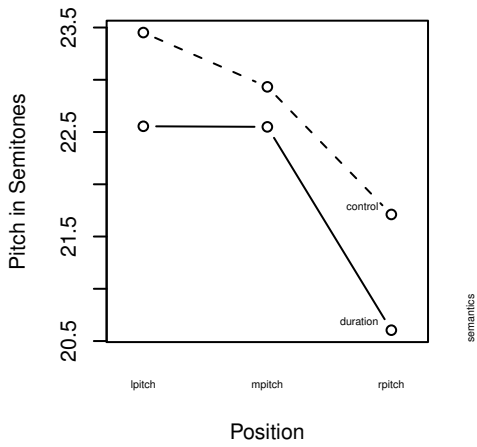


Figure 7.8: [*winter term*] + N3

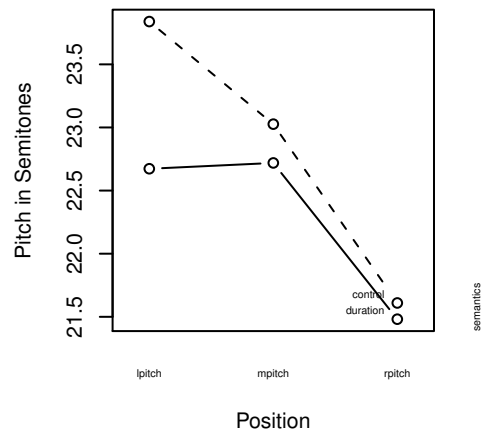


Figure 7.9: [*flight test*] + N3

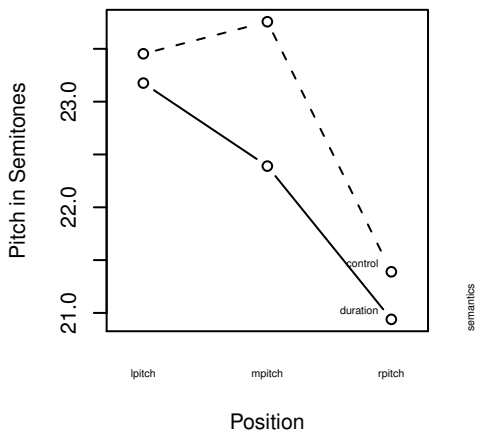


Figure 7.10: [*field trip*] + N3

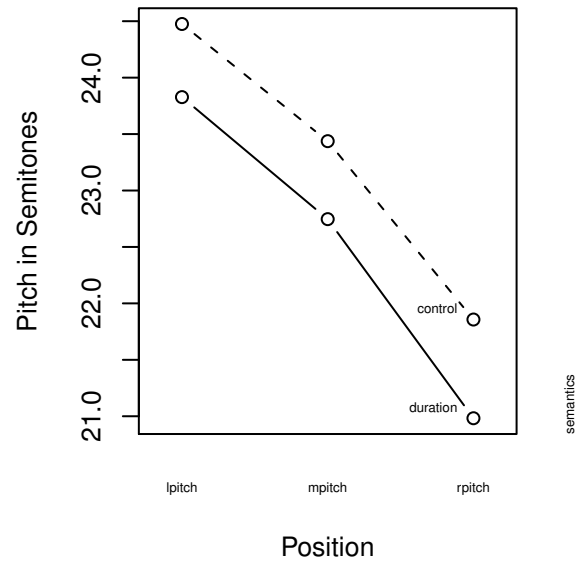


Figure 7.11: [*monday morning*] + N3

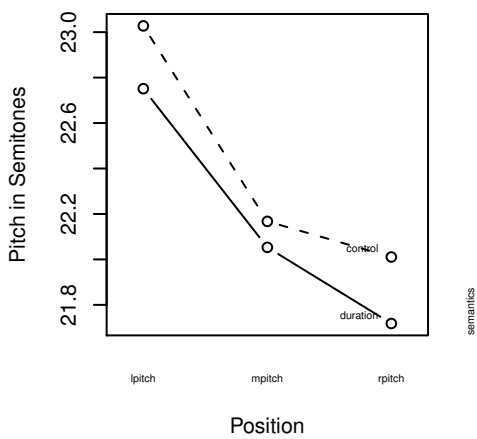
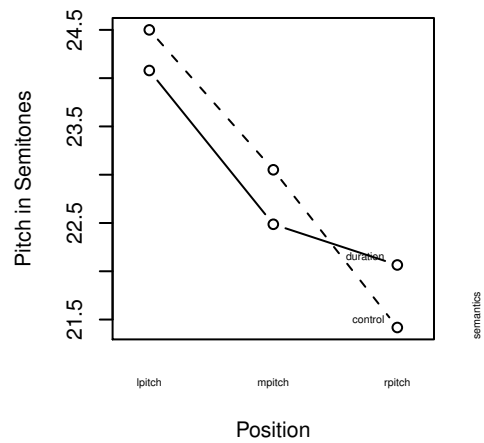


Figure 7.12: [*spring break*] + N3



and constituent N2 and the clearly lower pitch on constituent N3. In particular, the analysis revealed that pitch on constituent N1 in the test items of the *holiday season* subset is only 0.005 semitones higher than pitch on constituent N2. In the test items of the *winter term* compounds, constituent N1 has even a slightly (-0.005ST) lower pitch than constituent N2. This level pitch between constituent N1 and constituent N2 indicates right prominence between the two constituents, i.e. constituent N2 is more prominent than constituent N1 (e.g. Kunter 2011). Since constituent N3 has a clearly lower pitch in the two test items of the *holiday season* and *winter term* subsets, we may argue that constituent N2 is also the most prominent one. But what about the control items? The higher pitches on constituent N1 in the control items of the *winter term* and *holiday season* compounds may indicate that speakers assigned highest prominence to constituent N1 in these compounds. What may speak against this interpretation is the fact that the average pitch pattern of the control items differs from the one observed earlier for compounds with highest prominence on constituent N1; thus far left-branching compounds with highest prominence on constituent N1 exhibited a larger pitch drop between the first two constituents and a smaller pitch drop between constituent N2 and constituent N3 (see again the results sections of the corpus study and experiment 1). Yet, for the control items of the current two subsets, we observe a larger pitch change between constituent N2 and constituent N3 and a smaller pitch change between constituent N1 and constituent N2. This pitch behaviour rather resembles the one obtained for compounds with highest prominence on constituent N2 (e.g. R/N2 and L/N2 compounds in experiment 1). Thus, it may also be possible that the semantic and control items of these two subsets both have highest prominence assigned to constituent N2, yet that prominence on constituent N2 is more pronounced in the test items than in the control items.

Turning to the subset of the *flight test* compounds shown in Figure 7.9, we observe that the higher pitch on constituent N2 in the control items is due to a rising pitch of about 0.3 semitones between constituent N1 and constituent N2 in contrast to a falling pitch of about 0.8 semitones between these two constituents in the test items. According to the model of the *flight test* compounds this difference in the pitch change between constituent N1 and N2 ($p < 0.05$) is statistically significant, as is the difference between the means estimated for constituent N2 ($p < 0.001$) between the two groups. Apart from that, however, the control and test items do not differ significantly from each other in one of the two remaining pitches estimated by the model, i.e. in pitch assigned to constituent N1 and constituent N3, respectively. Thus, despite the same complex constituent, the semantic and control items differ in pitch assigned

to constituent N2. With reference to the compounds' prominence pattern, the result suggest that the control items exhibit highest prominence on constituent N2 and the test items of the *flight test* subset on constituent N1. Yet, it seems also plausible that both control and test items exhibit highest prominence on constituent N2, but that speakers pronounced this prominence pattern more clearly in the control items. This is indicated by the fact that for the control items we even observe a rising pitch on constituent N2. As previous studies by Kunter (2011) or Plag et al. (2008) dealing with prominence assignment to NN compounds have shown, highest prominence on constituent N2 is more often marked by equally high pitches on constituent N1 and constituent N2 or even a slightly lower pitch on constituent N2. Given this finding as well as the observation that the average pitch pattern of the test items of this subset is quite similar to the pitch pattern observed in previous models for compounds with highest prominence on constituent N2, it seems more likely that both in the test items and the control items of this subset constituent N2 is the most prominent one.

The fourth subset, which shows a difference in the pitch pattern of control and test items, is that of the *spring break* compounds shown in Figure 7.12. We observe that pitch drops to the same degree between constituent N1 and constituent N2 in both the test and control items ($p = 0.738$). Furthermore, we observe that the test items tend to have a higher pitch on constituent N3 than the control items, which is caused by a smaller drop in pitch between constituent N2 and constituent N3 in the test items. This pitch pattern is in line with the prediction of the IC-Prominence Hypothesis, which predicts the test items to have a higher pitch on constituent N3 than the control items. In fact, the model fitted to the *spring break* subset revealed that the pitch change between constituent N2 and constituent N3 is statistically significantly smaller in the test items than in the control items ($p < 0.05$). However, the test items and the control items do not significantly differ in the mean pitches estimated for constituent N3 ($p = 0.072$). This suggests that the pitch values on constituent N3 strongly overlap between control and test items. Hence, there is a tendency for speakers to assign a higher pitch to constituent N3 in the test items than in the control items of this subset, however, the difference is not large enough to be significant. Thus, contrary to the visual impression, there is no statistical support for the IC-Prominence Hypothesis.

Finally, a look at the two plots of the *monday morning* and *field trip* compounds reveals no difference between control and test items with respect to the degree to which pitch drops between any of the three constituents of the compounds. For both subsets we only observe that the control items tend to have higher pitches than the test items.

To sum up, the closer inspection of the individual items of the *IC2 is during IC1* com-

pounds showed that in three of the six subsets speakers assigned different pitches to the control and test items. For these three subsets the pitch differences between control and test items are found between the members of the complex constituents. Yet, as to whether these pitch differences between control and test items indicate different prominence patterns or whether one and the same prominence pattern is pronounced to a different degree by the speakers is difficult to say. Therefore, I rather remain agnostic with reference to an interpretation of the pitch patterns of these compounds in terms of prominence. Yet, we note that there are differences in pitch between control and test items, which may explain some of the high residuals observed earlier in this subsection in Figure 7.5. Furthermore, we note that none of the test items seems to behave as predicted by the IC-Prominence Hypothesis.

IC2 is located at IC1

Having investigated the group of *IC2 is during IC1* compounds in more detail, we now have a closer look at the *IC2 is located at IC1* compounds. In contrast to the *IC2 is during IC1* model, the model of the *IC2 is located at IC1* compounds did not reveal that speakers assigned generally higher pitches to the control items than to the test items. Because of that, we may not expect to find the same amount of variation between control and test items as observed for the *IC2 is during IC1* subset. Nonetheless, there may be some compounds in the data set that actually follow the predictions of the IC-Prominence Hypothesis.

Figures 7.13 - 7.18 show the six interaction plots for the six pairs of the *IC2 is located at IC1* compounds. In Figure 7.13 and Figure 7.14, we observe that the control and the test items of the two subsets behave exactly alike. They neither differ in the way pitch changes between each of the three constituents nor in the general pitch height.¹³ The pitch pattern of these compounds is similar to the average pitch pattern obtained for the *IC2 during IC1* subset. Furthermore, Figure 7.15 shows the interaction plot for the *gas station* subset. As for the *cocktail bar* and *movie theater* subsets, the test items and control items do not differ in the pitch relation between constituent N1 and constituent N2 nor in the pitch relation between constituent N2 and constituent N3. Yet, in contrast to the *cocktail bar* and *movie theater* subsets, the model of the *gas station* compounds revealed a significant main effect for the factor SEMANTICS. As shown in the plot speakers assigned in general significantly higher pitches to the control

¹³The respective models revealed no significant interaction of SEMANTICS by POSITION nor a main effect for the factor SEMANTICS.

Figure 7.13: [cocktail bar] + N3

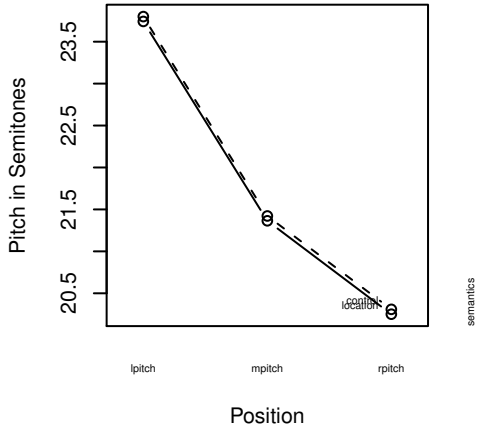


Figure 7.14: [movie theater] + N3

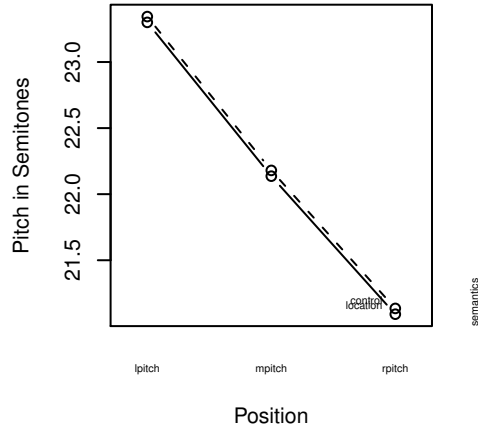


Figure 7.15: [gas station] + N3

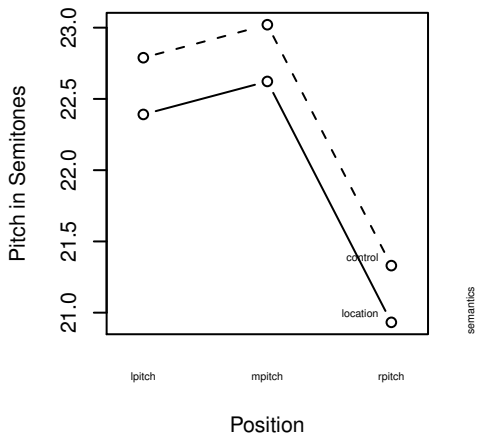


Figure 7.16: [coffee house] + N3

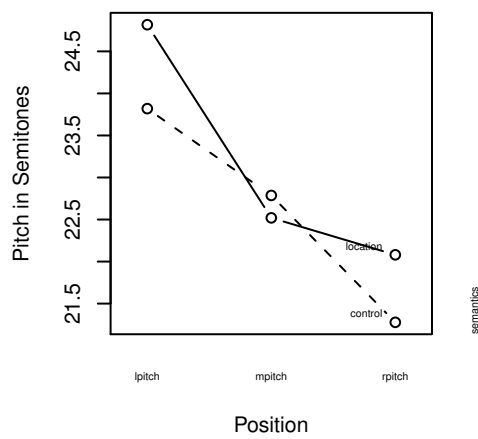


Figure 7.17: [apartment building] + N3

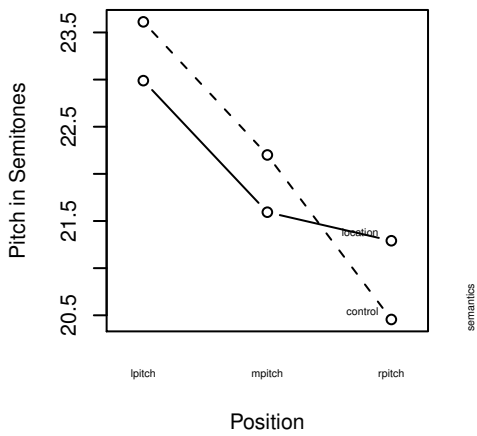
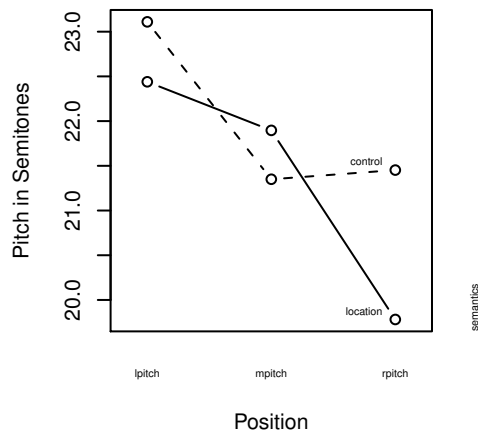


Figure 7.18: [hockey stadium] + N3



items than to the test items ($p < 0.05$).¹⁴ Furthermore, highest pitch is assigned to constituent N2 in the control and test items of the *gas station* subset, which strongly suggests that highest prominence is assigned to constituent N2 in these compounds.

In contrast to the three subsets discussed thus far, the remaining three interaction plots show some pitch variation between control and test items. In each plot, we observe that the test items differ from the control items in the degree to which pitch changes between at least two of the three constituents. Interestingly, in Figures 7.16 and 7.17, we observe that the test items tend to have a higher pitch on constituent N3 than the control items of the subsets, which points in the direction of the IC-Prominence Hypothesis. However, according to the model of the *coffee house* compounds neither the difference in the mean pitches of constituent N3 between the test and control items is statistically significant ($p = 0.0685$) nor the difference in the pitch drop between constituent N2 and constituent N3 ($p = 0.0882$). The model only reveals that the test items of the *coffee house* subset tend to have a higher pitch on constituent N1 than the control items ($p < 0.05$) as well as that pitch drops to a larger degree between constituent N1 and constituent N2 in the test items than in the control items of this subset ($p < 0.05$). The result provides no support for the IC-Prominence Hypothesis to be true for this subset. Yet, the different pitch patterns between control and test items suggest different prominence patterns: in the test items the most prominent constituent seems to be constituent N1, whereas in the control items it is constituent N2.

The situation is different for the compounds of the *apartment building* subset shown in Figure 7.17. As for the *coffee house* compounds, we also observe that test items have a higher mean pitch on constituent N3 than the control items and that the two groups differ in the degree to which pitch changes between constituent N2 and constituent N3. However, contrary to the *coffee house* subset, the model of the *apartment building* subset revealed that the pitch change between constituent N2 and constituent N3 is indeed significantly smaller in the test items than in the control items ($p < 0.01$). Furthermore, we find a significantly higher pitch on constituent N3 in the test items ($p < 0.05$). With reference to the two mean pitches of constituent N1 and constituent N2, the control and test items do not significantly differ from each other. Thus, the test items of the *apartment building* subset behave differently from the control items of the subset in that they tend to have a higher pitch on constituent N3. This result is in

¹⁴As for the *IC2 during IC1* compounds, the mixed-effects model of each individual item pair for which I determined statistically significant differences are documented in part B of the Appendix, whereas the mixed-effects models of item pairs for which I determined no significant interaction are not documented in the Appendix of this thesis.

line with the IC-Prominence Hypothesis.

Finally, the plot of the *hockey stadium* subset shows a totally different picture than the other plots. Contrary to the IC-Prominence Hypothesis, the control items of this subset have a higher pitch on constituent N3 (+1.67 ST) than the test items. According to the model this difference is also statistically significant ($p < 0.001$). Furthermore, we find that the pitch change between constituent N1 and constituent N2 is significantly larger in the control items than in the test items ($p < 0.05$). Yet, the two groups neither differ in pitch assigned to constituent N1 nor in pitch on constituent N2. The higher pitch on constituent N3 in the control items may suggest that some of these compounds exhibit highest prominence on constituent N3, contra the IC-Prominence Hypothesis. In contrast, the pitch pattern of the test items of this subset rather resembles that of compounds with highest prominence on constituent N2.

In summary, the inspection of the *IC2 is located at IC1* compounds revealed that for three of the six pairs the speakers participating in the experiment assigned different pitch patterns to the control and test items. The semantic and control items of the *coffee house* subset differ with reference to the pitches assigned to the members of the complex constituent. The pitch pattern for the *apartment building* compounds is in accordance with the IC-Prominence Hypothesis, i.e. the test items of the *apartment building* subset differ from their respective control items only in pitch assigned to constituent N3. However, the analysis also provides a counter-example to the IC-Prominence Hypothesis. For the *hockey stadium* subset, we find that the control items have a significantly higher pitch on constituent N3 than the regular test items.

7.3 Discussion

The experiment presented in this chapter investigated the predictions of the IC-Prominence Hypothesis according to which the same factors found to trigger right prominence in biconstituent compounds also trigger right prominence at the IC-level of left-branching triconstituent NNN compounds. These factors are semantics and analogy. The factor analogy was tested with the two lexical items *pie* and *avenue*, which were thought to cause left-branching compounds to exhibit highest prominence on constituent N3 in analogy to the head constituent. The factor semantics was tested for the two semantic relations *IC2 is located at IC1* and *IC2 is during IC1*. According to the IC-Prominence Hypothesis, compounds that can be interpreted in terms of one of those two semantic relations at the IC-level should have highest prominence on constituent N3. I will start the discussion with the result obtained for the factor analogy and particularly that of the *avenue* and *street* comparison before turning to *pie* compounds and the factor semantics.

7.3.1 Analogy

The pitch analysis revealed that *avenue* and *street* compounds had equally high pitches on constituent N1 but differed in pitch assigned to constituent N2 and constituent N3. To be more precise, *avenue* compounds had a higher pitch on constituent N2 and on constituent N3 than *street* compounds. The result that the two compound groups did not differ in pitch on constituent N1, but differed in pitch on constituent N3 is fully in accordance with the prediction of the IC-Prominence Hypothesis. It provides strong evidence for the assumption that the head constituent *avenue* causes speakers to assign highest prominence to constituent N3 in such compounds. In contrast to that, the pitch pattern determined for *street* compounds strongly suggests that highest prominence was assigned to constituent N1 in these compounds. This is indicated by the large pitch drop of about 2 ST between constituent N1 and constituent N2 and the smaller pitch drop of about 1 ST between constituent N2 and constituent N3. This pitch pattern of *street* compounds strongly resembles that of left-branching compounds with embedded left prominent NNs analysed in experiment 1 and that of the majority of left-branching compounds in the corpus study.

Yet, as noted in the previous paragraph, *avenue* compounds did not only exhibit a higher pitch on constituent N3, but also yielded a higher pitch on constituent N2 than *street* compounds. This result is quite unexpected given the way the compounds were constructed; the *avenue* and *street* compounds had the same complex constituents and

thus differed only in their head constituents from each other. Regarding this situation, we had rather expected that our speakers would drop their pitch to the same degree between constituent N1 and constituent N2 in *avenue* and *street* compounds and would only use different pitches on constituent N3. However, contrary to this assumption, the prominence difference between *avenue* and *street* compounds was marked by a difference in the pitch drop between constituent N1 and constituent N2 rather than between constituent N2 and constituent N3. The result indicates that in order to mark constituent N3 to be the most prominent one in *avenue* compounds, our speakers rather maintained a relatively high pitch on all three constituents of the *avenue* compounds instead of dropping and raising their pitch. This strategy seems quite in accordance with findings by Kunter (2011) and Plag et al. (2008), who showed that right prominence is more often marked by a level F0 contour rather than clearly falling and rising pitch contours. In contrast, assigning highest prominence to constituent N1 in *street* compounds does not require the speakers to maintain high pitches on constituent N2 and constituent N3. Instead, speakers may immediately drop their pitch after they have pronounced the first constituent. The result suggests that the prominence pattern of the complex constituent does not play a role when assigning highest prominence to the head constituent of left-branching compounds. Instead, as mentioned above, speakers seem to prefer the strategy of maintaining their pitch on a high level if they intend to mark the last constituent as being the most prominent one.

The result that we are indeed dealing with highest prominence on constituent N3 in *avenue* compounds but with left prominence in *street* compounds gains further weight if we consider intrinsic pitch differences. In the corpus study as well as in experiment 1 intrinsic pitch differences were neglected. First, because in the corpus study and experiment 1, vowels were randomly distributed and thus the effect of intrinsic pitch was thought to be largely neutralized. Second, because pitch was measured over the sonorant part of the rime rather than only the nucleus. Therefore, the measurement intervals often also included sonorants which were thought to reduce potential intrinsic pitch differences of the nucleus. Finally, the items were modelled as a random effect, which was also intended to account for at least part of the variance introduced by those intrinsic pitch differences.

With regard to the comparison of *avenue* and *street* compounds, however, none of these conditions was met. The sonorant part of the most prominent rime of the two head constituents in *avenue* and *street* compounds happened to consist only of a nucleus. In particular, in the analysis we constantly compared the high front vowel [i] of *street* with the low front vowel [æ] of *avenue*. As already mentioned in section 3 high

vowels exhibit higher intrinsic pitches than low vowels (cf. Whalen and Levitt 1995; Neppert and Pétursson 1986). Hence, *avenue* should have a lower pitch than *street*, if prominence does not play a role. Given this lower intrinsic pitch of *avenue*, the result that *avenue* compounds have a higher pitch on constituent N3 than *street* compounds strongly supports the interpretation that speakers marked the head of *avenue* compounds as more prominent than the head of *street* compounds.

Finally, the pitch patterns obtained for the *avenue* and *street* compounds may also be phonologically interpreted in terms of pitch accents. The pitch pattern of *street* compounds strongly suggests that there is a single pitch accent assigned to constituent N1 with no further pitch accents assigned to constituents N2 and N3. In contrast to that, the pitch pattern of the *avenue* compounds indicates that we are dealing with three pitch accents and in particular three downstepped pitch accents as indicated by the constant pitch drop between all three members of the *avenue* compounds. In a row of such downstepped pitch accents, it is assumed that the last of these accents is usually perceived as the highest prominent one.¹⁵

Turning to the comparison of *pie* and argument-head compounds, the analysis revealed that *pie* compounds had a significantly higher pitch on constituent N2 than argument-head compounds. For the pitches on constituent N1 and constituent N3, the analysis revealed no difference between *pie* and argument-head compounds. Hence, the analysis provides no evidence for the lexical item *pie* to trigger highest prominence on the head constituent of triconstituent left-branching compounds. Instead, the result suggests that *pie* compounds have highest prominence on constituent N2 and argument-head compounds on constituent N1. Given that *pie* and argument-head compounds had the same complex constituents, this result is surprising. In the absence of the predicted effect of the IC-Prominence Hypothesis, we would have rather expected *pie* and argument-head compounds to behave exactly alike. In particular, both should have either exhibited highest prominence on constituent N1 or on constituent N2. In contrast, in the presence of the analogical effect, *pie* compounds should have differed from argument-head compounds in the mean pitch assigned to constituent N3, or in pitch on constituent N2 and constituent N3 as previously observed for *avenue* and *street* compounds. Thus, we are left with the question of why we find this difference between *pie* and argument-head compounds.

It may be possible that the pitch behaviour of *pie* compounds is nevertheless triggered by analogy. Yet, the analogical pattern triggered by *pie* as head constituent is

¹⁵A constant pitch drop between pitch accents is argued to be characteristic of downstepped pitch accents, which contrast with non-downstepped accents (Ladd 1996).

not that of highest prominence on constituent N3, as predicted by the IC-Prominence Hypothesis, but instead the analogical pattern triggered by *pie* is that of highest prominence on constituent N2. Thus, the lexical item *pie* may not itself attract prominence, but instead triconstituent *pie* compounds may in general exhibit highest prominence on constituent N2. This would explain why the *pie* and control items differed from each other, although they contained the same complex constituents. This assumption, however, would also suggest that highest prominence on constituent N2 in left-branching compounds is not only triggered by embedded right prominent NN compounds as shown in experiment 1, but that the third constituent may also trigger this prominence behaviour. Yet, this hypothesis clearly requires more empirical research.

To sum up, the analysis of the *avenue* and *street* data provides strong evidence that in left-branching compounds with *avenue* as head constituent highest prominence is assigned to constituent N3, whereas in *street* compounds highest prominence is assigned to constituent N1. The result provides evidence that the well-known *street* vs. *avenue* contrast observed for biconstituent compounds is also prevalent in triconstituent NNN compounds. The result supports an analogical account towards prominence assignment in triconstituent compounds. In contrast, the analysis of the *pie* compounds provided no evidence for the assumption that the lexical item *pie* triggers highest prominence on constituent N3 in left-branching compounds. Instead, the result suggests that speakers assigned highest prominence to constituent N2 in *pie* compounds and to constituent N1 in the control items. The result does not support the factor analogy as predicted by the IC-Prominence Hypothesis. Still, *pie* compounds exhibited a different prominence pattern than the corresponding control items, which shows that the choice of *pie* as head constituent influenced the speakers' prominence assignment. Yet, as to whether analogy is responsible for this trend observed for *pie* compounds requires some further research.

7.3.2 Semantics

The statistical analysis of the two semantic relations *IC2 during IC1* and *IC2 is located at IC1* provided no evidence in favour of the IC-Prominence Hypothesis. Contrary to the hypothesis, the test items of the *IC2 is located at IC1* and *IC2 during IC1* subsets did not systematically differ from their respective control items in pitch on constituent N3 and thus in prominence assigned to constituent N3. Hence, according to the present data neither the semantic relation *IC2 during IC1* nor *IC2 is located at IC1* systematically triggered right prominence at the IC-level of left-branching NNN compounds. Instead,

the result of the pitch analysis indicates that the semantic and control items of the *IC2 during IC1* and *IC2 is located at IC1* subsets exhibit highest prominence on one of the two members of the complex constituents. In particular, for the semantic and control items of the *IC2 during IC1* compounds, we observed that pitch dropped on average half as much between constituent N1 and constituent N2 than between constituent N2 and constituent N3, a pitch pattern similar to that observed for left- and right-branching compounds with highest prominence on constituent N2 in experiment 1. In contrast, for the control and test items of the *IC2 is located at IC1* compounds, we observed a generally larger pitch drop between constituent N1 and constituent N2 than between constituent N2 and constituent N3. This pitch pattern of the *IC2 is located at IC1* compounds and their respective control items rather resembles that of left-branching compounds with highest prominence on constituent N1 as observed in experiment 1 and in the corpus study. Hence, the pitch patterns of the two semantic groups suggest that for the majority of the *IC2 during IC1* compounds the most prominent constituent is constituent N2, whereas for the majority of the *IC2 is located at IC1* compounds highest prominence is assigned to constituent N1.

The seemingly different pitch patterns of *IC2 during IC1* compounds and *IC2 is located at IC1* compounds may be explained by the fact that I have not controlled for the prominence pattern of the embedded NN compounds. As argued in subsection 7.1.1, controlling for the prominence pattern of the embedded NN was not considered to be relevant for testing the IC-Prominence Hypothesis, as irrespective of whether the complex constituent is left or right prominent, the control and test items should differ in pitch assigned to constituent N3. Thus, it may be possible that the embedded NNs of the *IC2 during IC1* subset were primarily right prominent whereas those selected for the *IC2 is located at IC1* subset were generally left prominent.

Apart from these general trends determined for the two semantic relations, however, a closer inspection of the 12 individual item pairs also revealed one item that behaved as predicted by the IC-Prominence Hypothesis. The item *apartment building party* tended to have a higher pitch on constituent N3 than its corresponding control item *apartment building manager*. However, it remains questionable whether the right prominence of this individual compound is indeed due to the semantic relation *IC2 is located at IC1* since this item was the only one in the data set that behaved according to the hypothesis. Thus, highest prominence on constituent N3 in the *apartment building* compounds may have also been triggered by other factors than the semantic relation of *IC2 is located at IC1*. One such factor may be analogy, for which the experiment in fact provided some empirical support. This assumption is further supported by the

fact that the analysis also revealed a counter-example to the IC-Prominence Hypothesis, namely that of *hockey stadium cleaner*. This control item had a significantly higher pitch on constituent N3 than the test item *hockey stadium event*.

In summary, the investigation of the two semantic relations *IC2 is during IC1* and *IC2 is located at IC1* provided no evidence for the correctness of the IC-Prominence Hypothesis with reference to the factor semantics. Thus, neither the *IC2 is during IC1* nor the *IC2 is located at IC1* relation generally triggered highest prominence on constituent N3.

7.4 Conclusion

The aim of the present experiment was to test whether highest prominence on constituent N3 in left-branching NNN compounds is triggered by analogy to the head constituent on the one hand, and specific semantic relations at the IC-level of the compounds on the other hand. The pitch analysis showed that the *avenue* and *street* contrast observed for biconstituent Noun+Noun compounds (e.g. Ladd 1984; Liberman and Sproat 1992) also exists for triconstituent Noun+Noun+Noun compounds. The result obtained for the *avenue* and *street* compounds provides evidence for the assumption that analogical mechanisms also play a role in prominence assignment to left-branching NNN compounds. Yet, the empirical support is restricted to the *avenue* and *street* compounds; compounds with *pie* as head constituent did not differ from their control items as predicted by the IC-Prominence Hypothesis, but instead *pie* compounds differed from their control items with reference to pitch assigned to constituent N2. The result implies that speakers assigned highest prominence to constituent N2 in *pie* compounds but to constituent N1 in the control items. As to whether this result obtained for the *pie* compounds is also triggered by analogy, as suggested in subsection 7.3.1, however, cannot be answered here and thus remains subject to future research. For instance, one may replicate the present experiment with different, and perhaps more *pie* compounds, as the ones analysed in the current experiment in order to find out if *pie* compounds exhibit the same pitch pattern as they did in this experiment. This should be the case if the choice of *pie* as head constituent generally causes triconstituent compounds to have highest prominence on constituent N2.

Turning to the factor semantics, we found that neither of the two semantic relations *IC2 during IC1* and *IC2 is located at IC1* generally triggered highest prominence on constituent N3 in left-branching compounds. Hence, we may conclude that the semantic relation at the IC-level of left-branching compounds does not influence prominence

assignment in NNN compounds.

In general, the result of the present experiment is in line with findings of studies dealing with prominence assignment in NN compounds (Lappe and Plag 2007; Plag et al. 2008; Plag 2010). These studies found that prominence variation in NN compounds is in general better explained by predictors associated with the factor analogy than by the factor semantics. As noted earlier in section 2.1, Plag (2010:272) showed that the so-called constituent family bias of a given compound constituent is a stronger predictor with reference to prominence assignment in NN compounds than the semantic relations proposed in the literature. The experiment presented in this chapter also provides evidence for the influence of analogy but no empirical support for the assumption that the semantic relation between the compound constituents influences prominence assignment in triconstituent compounds.

8 Summary and conclusion

This thesis presented an empirical investigation of the prominence behaviour of English triconstituent NNN compounds by analysing a large number of speech corpus and experimental data. It is the first study that tested Liberman and Prince's (1977) generalizations regarding prominence assignment in triconstituent compounds by using acoustic measurements of pitch in order to determine the compounds' prominence behaviour. Furthermore, I proposed and tested two alternative hypotheses to the LCPR, which also incorporated the existence of right prominent NNs into their predictions. This chapter summarizes the main findings of the thesis, provides some concluding thoughts and makes suggestions for future research in this area.

According to the Lexical Category Prominence Rule proposed by Liberman and Prince (1977), prominence assignment in triconstituent NNN compounds depends on the branching direction of the compound; left-branching compounds have highest prominence on constituent N1 whereas in right-branching compounds constituent N2 is marked as the most prominent constituent of the whole compound. The corpus study presented in chapter 4 is the first of its kind that tested these generalizations regarding prominence assignment in triconstituent NNN compounds in English. The statistical analysis of the corpus data revealed that the majority of left- and right-branching NNN compounds tend to behave as predicted by the Lexical Category Prominence Rule. Yet, the study also showed that the rule is far from being categorical in nature; about 36% of the left-branching compounds and 30% of the right-branching compounds of the corpus data exhibited a prominence pattern contrary to the rule. In particular, I also found left- and right-branching compounds with highest prominence on the right member of the complex constituent (e.g. *[science fiction] shocker*, *state [health program]*) and on the constituent outside the complex constituent (e.g. *China [information center]*, *[child care] crisis*).

This large proportion of violations to the rule is not surprising, however, given the theoretical shortcomings associated with the LCPR on the one hand, and findings of the few previous studies dealing with prominence assignment in NNN constructions on the other hand (e.g. Kvam 1990; Giegerich 2009). The result of the corpus

study provides strong empirical evidence against branching direction as the governing force in prominence assignment to NNN compounds and strongly suggests that their prominence pattern must also depend on factors other than the branching direction.

One such factor is the prominence pattern of the embedded NN compound as shown in the production experiment presented in chapter 6. The production experiment focused on the question whether embedded right prominent NN compounds affect the prominence behaviour of triconstituent NNN compounds in that left-branching compounds have highest prominence on constituent N2 and right-branching compounds on constituent N3, contra the LCPR. In addition, it also addressed the question whether compounds with embedded left prominent NNs, i.e. compounds meeting the underlying assumptions of the LCPR, behave as predicted by the rule. The two predictions were subsumed under the heading of the Embedded Prominence Hypothesis (EPH).

The experiment revealed a general trend towards the EPH for left- and right-branching compounds with embedded left prominent NNs ('L/N1'; 'R/N2') as well as for left-branching compounds with embedded right prominent NNs ('L/N2'). However, the EPH was not supported for right-branching compounds with embedded right prominent NNs ('R/N3'). This result obtained for R/N3 compounds is rather surprising given that the group of L/N2 compounds behaved as predicted by the EPH and we would have rather expected the EPH to be true for either both compound types or none of the two. Yet, as discussed in section 6.3 it may be possible that the lack of empirical support for the EPH and the group of R/N3 compounds is due to unknown factors operating independently at the IC-level of triconstituent compounds. These factors may have obscured the predicted effect of the embedded right prominent NNs in that they triggered highest prominence on constituent N1 in some of the R/N3 compounds.

What follows from this experiment is that the EPH makes the correct predictions for left- and right-branching compounds with embedded left prominent NNs. Furthermore, despite the lack of empirical support for the EPH in the group of R/N3 compounds, the result obtained for L/N2 compounds provides first empirical evidence for the assumption that the existence of right prominent NNs directly influences prominence assignment in triconstituent compounds. As to whether the EPH and its predictions for embedded right prominent NNs are indeed generally restricted to left-branching compounds, or whether the absence of an effect in R/N3 compounds is peculiar to the present data set, however, requires further empirical research.

Apart from investigating the influence of embedded left and right prominent NN compounds on the prominence behaviour of left- and right-branching NNN compounds, the present thesis also had a closer look at two factors potentially responsible for right prominence at the IC-level of left-branching NNN compounds. Thus far highest prominence on constituent N3 in left-branching NNNs has been justified by simply assigning a phrasal status to these constructions (e.g. Liberman and Prince 1977; Sproat 1994). Yet, as noted before, a distinction between NNN compounds and NNN phrases solely based on their prominence pattern is questionable due to the lack of other convincing criteria to differentiate between the two types of constructions (e.g. Bell 2011). The IC-Prominence Hypothesis (IPH) formulated in chapter 2.2 presents an alternative approach to the phrasal analysis of left-branching NNNs with highest prominence on constituent N3. The IPH predicts that the same factors triggering right prominence in NN compounds also trigger right prominence at the IC-level of left-branching NNN compounds, contra the LCPR. The two factors tested in this thesis were that of semantics and analogy, both of which have been found to account for some of the variation observed for NN compounds (e.g. Lappe and Plag 2007; Plag et al. 2008; Plag 2010).

The IC-Prominence Hypothesis was tested in a second production experiment described in chapter 7. The experiment provides some empirical support for the IC-Prominence Hypothesis and the factor analogy, but no support for the factor semantics. As predicted by the IPH left-branching compounds with *avenue* as head constituent tended to have highest prominence on constituent N3, whereas the same compounds ending in *street* showed a tendency towards highest prominence on constituent N1. The result provides evidence that the same *avenue* and *street* contrast observed for biconstituent compounds also exists for triconstituent compounds. In more general terms this result implies that the same analogical mechanisms governing prominence in NN compounds must also operate at the IC-level of triconstituent compounds. Despite this outcome, however, we must also note that the effect of analogy is only found for the *avenue* and *street* compounds in this thesis, as left-branching compounds with *pie* as head constituent did not behave as predicted by the IPH. In the light of this result, we conclude that the current thesis indeed provides first empirical proof towards the assumption that prominence assignment in triconstituent compounds works on the basis of analogy, yet more research with more data seems necessary to further substantiate this claim.

For instance, one could conduct a corpus study along the lines of Plag's (2010) study and try to calculate constituent family biases for the three members of triconstituent

compounds. As noted earlier in this thesis, the constituent family bias, which provides a measure for the tendency of a given constituent family to favour a particular prominence pattern, turned out to be a robust predictor for prominence assignment in biconstituent compounds (cf. Plag 2010).

Crucially, such a study should also include the analysis of right-branching compounds, for which the two factors analogy and semantics were not explicitly tested in the present thesis. Yet, if prominence assignment in triconstituent compounds in fact works on the basis of analogy as suggested by the *avenue vs. street* result, influences of analogy may also account for potential IC-level violations in right-branching compounds. Besides, future studies dealing with prominence assignment in triconstituent compounds should also incorporate those informativeness measures which turned out to be significant in the study by Bell and Plag (2012). As noted in section 2.1, the influence of informativeness was not tested in this thesis; however, if prominence assignment in triconstituent compounds is indeed governed by the same factors as those operating in biconstituent compounds, effects of informativeness are also highly expected for triconstituent compounds. In fact, some first hints for such an influence of informativeness are provided in the corpus study of this thesis (see subsection 4.3.2 again).

The results obtained in this thesis, however, do not only seem to have consequences for triconstituent left- and right-branching compounds; but the failing of the Lexical Category Prominence Rule has also serious repercussions for prominence assignment in structurally ambiguous compounds as well as for larger compounds. As already mentioned earlier in this thesis, it is generally assumed that the prominence pattern assigned to structurally ambiguous compounds functions as a means to disambiguate such compounds (e.g. Liberman and Prince 1977; Selkirk 1984; Liberman and Sproat 1992; Cinque 1993; Visch 1999). For instance, the compound *kitchen towel rack* is said to be structurally ambiguous as it may be interpreted as either a 'rack for a kitchen towel' (left-branching) or a 'towel rack located in the kitchen' (right-branching). Depending on its interpretation, speakers assign different prominence patterns to the compound. In particular, in case of a left-branching interpretation the compound is claimed to have highest prominence on constituent N1, whereas a right-branching interpretation is assumed to be indicated by the assignment of highest prominence to the second constituent of the entire compound. However, given the result of experiment 1 that in both left- and right-branching compounds constituent N2 may be the most prominent constituent, the assumption that prominence has the function to disambiguate a structurally ambiguous compound becomes highly questionable.

Furthermore, the prominence pattern of compounds with four or five constituents is assumed to be governed by the same rule that governs prominence assignment in triconstituent compounds, namely the LCPR and the factor branching direction (e.g. Liberman and Sproat 1992:147). Hence, given that the LCPR fails to account for the prominence pattern of triconstituent NNN compounds, it is reasonable to believe that this is also the case for larger compounds.

In addition, in this thesis I measured the fundamental frequency, i.e. pitch, as an acoustic cue to prominence in order to determine prominence differences between different types of compounds. The F0 has been found to be one of the major acoustic cues of prominence in compounds. Furthermore, Kunter (2011) found empirical evidence that pitch cues the presence or absence of pitch accents in compounds. Hence, the pitch patterns obtained for the different compounds may also be phonologically interpreted in terms of pitch accents. According to this thesis, left-branching compounds with highest prominence on constituent N1 seem to have a single pitch accent on N1 but no pitch accent on constituent N2 and constituent N3. In contrast, left- and right-branching compounds in which constituent N2 is marked as the most prominent constituent typically have a pitch accent on the first and second constituent of the compound, but no accent on the third constituent (e.g. L/N2 and R/N2 compounds). Thus, the difference between compounds with highest prominence on constituent N1 and compounds with highest prominence on constituent N2 seems to be marked by the absence or presence of a pitch accent on the second constituent of a triconstituent compound. Finally, triconstituent compounds with highest prominence on constituent N3 seem to have a pitch accent on all three compound constituents, irrespective of their branching direction. This is indicated by the pitch patterns observed for the left-branching *avenue* compounds and the right-branching R/N3 compounds of the excluded speaker. For both groups, we observed a constant pitch drop of only 1 semitone between each of the three compound constituents. This constant downstep of pitch is characteristic of so-called downstepped pitch accents (cf. eg. Ladd 1996:77). Thus, the relatively high pitch values on all three compound constituents in these compounds may suggest a sequence of three downstepped pitch accents, with the last of the three accents being the most prominent one.

In general terms, the current thesis has shown that prominence assignment in triconstituent NNN compounds is more variable than generally assumed and that the factor branching direction fails to account for the prominence patterns of a large number of compounds. The result of the corpus study and the two experiments strongly suggests that any approach that tries to account for prominence assignment in left-

and right-branching triconstituent NNN compounds also needs to incorporate the existence of right prominent NN compounds.

A Data

Corpus study

The following list contains the 326 left-branching compounds that were extracted from the *Boston University Radio Speech Corpus* and which entered the final pitch analysis. For each compound the list also provides the name of the audio file from which the compound was extracted.

abortion rights advocate F2BS31P2
abortion rights challenger M4BS18P1
abortion rights groups F1AS39P4
abortion rights petition F2BS01P5
age bias complaint F3ASQ8P1
aids action committee F2BS21P3
airline industry F3ASV3P1
art student seminars M2BS13P8
auto body association F2BS27P2
auto emissions standard F1AS24P6
auto insurance reform F2BS27P1
ballot initiative revolution F3AS28P1
bank fund reorganization M4BS61P1
bank teller machine M2BS22P7
banking industry experts M2BS32P8
bar room fight M3BS05P9
base-ball owners F3AS59P1
baseball fans F3ASR6P1
baseball talk F3AS38P1
bass ale firm M2BS23PA
bathroom taps M2BS01P8
bay state drivers F3ASF5P1
bay state hospitals F1AS17P3
bay state republicans F3ASE7P1
bay state residents M1BS10P2
bay state voters M1BS24P2
beacon hill committee F3ASV9P1
beacon hill democrats M4BS37P1
beacon hill insiders M1BS11P1
beauty salon business F3ASL1P1
belmont holding pond M2BS18PA
birth control devices F2BTRLP1
birth control pills F1ATRLP3
birth control system F3ASC2P1
birthday cake F2BS12P2
blockbuster movies F3ASL2P1
blood-alcohol content F1AS20P5
blood-alcohol level F1AS21P6
blood lead levels F2BS02P4
bone marrow cells M3BS10P5
bone marrow donor M3BS12P1
bone marrow registries F3ASJ9P1
boston area artists F2BS22P1
boston area communities F3AS20P1
boston area galleries F3ASM3P1
boston area gangs F3ASV8P1
boston area hospitals F3AS92P1
boston area residents M3BS01P4
boston business magazine F3AS75P2
boston globe articles F3AS42P1
boston globe report M1BS27P4
boston globe reporter M1BS27P4
boston herald poll M4BS46P4
boston phoenix reporter F2BS16P1
boston university president F3ASV5P1
boston university school F1ARRLP3
brainstorming session F3ASU0P1
budget balancing measures M4BS09P6

budget balancing plan M4BS05P1
budget balancing targets M4BS09P5
budget cutting measures F3AS19P1
budget cutting plan M4BS01P3
budget cutting proposals M4BS55P2
bush administration officials M4BS10P5
cambridge-boston border F3AS72P1
cambridge hospital intern F2BS24P1
cambridge school committee F3AS60P1
cancer presumption bill M4BS13P3
capital gains levy M4BS12P4
capital gains tax F3AS06P1
catbird seat M2BS29P7
charles river esplanade F3ASA7P1
check kiting scheme M2BS14P6
chestnut hill M2BS13P4
child care crisis F1AS15P4
child care money F1AS15P4
child custody case F1AS32P4
child neglect law M1BS12P3
child neglect quote M4BS32P5
child neglect statute M1BS12P2
child support collections F1AS35P4
child support payments F1AS35P3
childbirth class F2BS20P1
chronicle publishing company M1BS37PB
city planning director M4BS60P5
classroom participation M1BS01P4
cleveland clinic foundation M3BS21P5
cloverleaf proposal M4BS19P6
coca growers associations M2BS34PF
coca leaf production M2BS34P4
color video monitors M3BS07P2
compound rate hike M2BS10P3
computer chip business M2BS35P2
computer science professor M1BS31P1
computer tabulating systems M2BS22P1
conservation law foundation F3ASU5P1
convention center business M4BS61P1
copley plaza hotel M1BS17P7
cost saving measures F1AS36P3
cost saving plan M4BS05P2
county jail inmates F1APRLP4
court room experience F3ASM0P5
courtroom theater F3ASM0P7
credit availability problem M1BS33P3
credit card approval M2BS06P4
credit card companies F3ASL3P1
credit card industry F3ASL3P1
credit card shuffle F3ASL3P1
curbside recycling M1BS16P1
day care centers F3ASJ3P1
day care provisions F1AS15P5
day care space F1AS15P4
defense department money M2BS03P2
defense industry lobbyist M4BS36P6
deficit reduction bill M4BS49P2
deficit reduction plan M4BS14P1
deposit insurance company F3ASP9P1
deposit insurance fund M4BS61P1
disaster relief bill F3AS27P1
district court judge F3ASE9P1
drug control policy F3AS15P1
drug control unit F2BS06P3
drug fighting efforts F1AS29P3
drug prevention policy M4BS10P3
dukakis administration officials F1AS43P4
dukakis administration plan F1AS29P3
dukakis administration resignees F1AS42P4
elizabeth island company M2BS08P4
emergency relief aid F3AS27P1
emergency room doctors M2BS25P7
ethics committee investigation F2BS08P6
examination division chief M2BS31PC
excise tax liabilities M2BS07P1
family estate planning M2BS12P2
family life education F1ATRLP3
farm labor camps F2BS05P2
felony sodomy charges F1AS35P2
flagship program M1BS31P3
food industry section M3BS03P2
football games M4BS05P1
football team M1BS38P1
franklin park zoo F3ASE2P1
gay rights activists M1BS29P1
gay rights bill F1AS23P5
gay rights group M1BS30P4
gay rights legislation F3AS24P1
governor sergeant appointee F1AJRLP6
grass roots advocates M4BS04P2
hair focus salon F3ASL1P1
hand gun control M4BS23P5
harness racing track F1AS28P6
health care advocates F3AS05P1
health care committee F2BS24P1
health care facilities M4BS16P4

health care services F3ASJ7P1
health care system M4BS29P3
health center director F2BS11P1
health center officials M4BS29P4
health insurance coverage F3AS33P1
health insurance issue F1AS37P6
health insurance plans M3BS16P2
healthcare law M4BS02P2
healthcare program M4BS04P4
healthcare proponents M4BS02P5
healthcare quality F2BS05P3
heartbreak hill F2BS15P1
home care products M2BS35P4
home improvement companies M2BS32P1
home improvement firms M2BS32P8
home improvement loans M2BS32P4
home improvement scam M2BS32P1
hometown politics M4BS57P2
homework assignment F2BS18P1
honeymoon period M4BS35P4
household average M3BS01P4
housekeeping errors F1AS42P6
hyde park section M1BS12P1
income tax deduction M2BS12P3
income tax hike M1BS11P4
income tax rate M1BS10P1
insurance funds collapse M1BS34P7
job elimination approach M4BS01P3
job training class M4BS09P1
job training program M4BS09P1
johnson controls employees F2BS02P1
junk bond status M4BS55P2
justice department probe F3AS11P1
justice system bureaucracies M1BS35P4
kansas city batters M1BS28P3
kansas city royals F1AS15P7
key note speaker M1BS14P3
law enforcement experts M4BS38P1
law enforcement officials F3ASU4P1
league championships series F3ASM9P1
letter writing campaign M1BS14P4
lewis-harris poll F2BTRLP1
line item veto M4BS30P3
litmus test supporters M4BS33P3
loan shark scam M2BS32P7
locker room vestibule M2BS13P1
london stage musical F3AS42P1
lotus development corporation F1AS31P4
madison park appearance M1BS17P3
manslaughter trial F3AS95P1
marrow-donor searches M3BS12P8
mattapan roxbury area F2BS16P3
menu printing business M2BS27P7
metro park zoos M1BS22P3
minority home owners M2BS33P1
mission hill area F3AS31P1
mission hill neighborhood F3AS31P1
mission hill resident F2BS20P3
mission hill shooting F3AS31P1
money generating measures M4BS05P1
money saving measures M4BS05P3
mortgage loan forms M2BS33P1
motor vehicle homicide F2BS31P1
motorcycle accident F2BS15P3
mount auburn club M2BS13P1
mount washington hotel F3ASX9P1
news conference luncheon M2BS26P1
newspaper ads M2BS33P3
newspaper business M1BS37P9
newspaper stories M1BS27P4
newsweek magazine F3AS18P1
newton-wellesley hospital M1BS25P1
nighttime fun F3AS83P1
nursing home care M4BS16P2
nursing home patients M4BS16P4
nursing home program M4BS16P1
oakland athletics F3ASF0P1
oil company executive F2BS12P4
oil ministry officials F1AS39P2
parking tax proposal M4BS60P6
pay cut proposal M4BS53P1
pay raise battle M4BS53P1
pay raise feud M4BS53P4
payroll tax M4BS02P1
peace corps program F3ASX0P1
piece-meal solutions M4BS01P1
plasma fusion center M3BS15P4
pork barrel politics M4BS61P1
project hope shelter M4BS07P6
property tax increases F3AST0P1
property tax limit F3ASU8P1
providence business news M1BS34P3
quality control chemist M2BS18P8
quincy-roxbury compromise M1BS20P1
rate payer rights M2BS10P1
rental assistance program M4BS07P8

rental housing association M2BS24P3
retail sales transactions M2BS06P4
riot management exercises F1AS28P5
road construction companies M2BS02PA
roadside test F3AS81P1
rockport artist F2BS33P1
rockport residents F2BS33P1
saint regis reservation F2BS18P1
sales tax base M1BS10P2
sales tax measure M4BS26P1
school committee meeting F2BS11P3
school committee members F2BS11P3
school committee restructuring F2BS10P3
science fiction shocker F3ASJ3PB
seabrook electricity M2BS10P6
seabrook opponents M2BS10P1
seabrook plant M3BS06P2
seat belt law F2BRRLP5
sex practice studies M2BS20P9
sheet metal plates M1BS28P4
shoestring budgets M3BS17P6
soul music circuit F3AS65P1
springfield jail F3APRLP4
stanley cup finals F3ASN7P1
state budget cuts F1AS43P5
state budget matters M4BS61P1
state house lobbyists M4BS43P2
state house observer F3ASJ4P1
state house reporter F3ASD3P1
state house sources M4BS45P2
state lottery officials M4BS05P1
stockbroker job F2BS23P1
street lawyers class M1BS06P8
street lawyers program M1BS06P3
street lawyers seminars M1BS06P7
substance abuse programs M4BS10P2
suffolk county prosecutor M1BS27P4
summer jobs programs M4BS20P3
task force member F1ATRLP3
task force report M1BS35P4
tax amnesty program F3ASP8P1
tax cutting petition M4BS43P6
tax equity alliance M1BS10P2
tax payers foundation M4BS12P3
tax preparation program M1BS07PB
textbook covers F3ASL6P1
thanksgiving day F3ASI6P1
tissue-type match M3BS12P2
traffic management plans M4BS56P1
transition team member M4BS35P2
triad design exhibit M2BS01P2
tufts university president F2BS29P3
tufts university researcher M3BS04P6
union army uniforms M2BS03P1
urban design specialist M4BS19P5
wall street firm M4BS26P3
wall street journal F3AST1P1
warner insurance company F1AS13P6
warning notice rules M2BS16P6
waste company officials M4BS06P1
waste disposal business M4BS06P1
water emergency rules F1AS40P5
water quality chemist M2BS18P1
water resources authority M4BS41P2
water treatment chemistry M2BS18P7
weekend nights F3AS83P1
weekend series F3ASN6P1
weekend warriors F3ASJ3PF
westfield maker M2BS35P4
westfield republican F3ASD4P1
wheelchair division F2BS15P3
wheelchair marathoners F2BS15P2
wheelchair racers F2BS15P2
wheelchair racing F2BS15P4
worcester foundation scientists M2BS15P4
world health organization F2BS29P1

The following list contains the 122 right-branching compounds that were extracted from the *Boston University Radio Speech Corpus* and which entered the final pitch analysis. For each compound the list also provides the name of the audio file from which the compound was extracted.

area street maps M2BS32P3
arlington town manager M4BS12P1
army family counselling F3ASJ3PC
beatrice foods company F3AS36P1
boston aids consortium M2BS20P8
boston city council F2BS03P3
boston city counselor F2BS07P4
boston district office M2BS31P7
boston finance commission F2BS10P4
boston gang members F2BS16P1
boston newspapers F3AS85P1
boston school committee F1ATRLP7
boston school superintendent F1AS34P4
boston state representative M4BS25P1
brick townhouses M3BS08P8
cabinet pay raise M4BS53P3
cambridge headquarters M1BS31P1
cambridge shoemaker M2BS35P4
cambridge water supply M2BS18P4
china information center M1BS02P6
community investment coalition M2BS33P5
community lending practices M2BS33P3
community meeting hall F2BS33P2
company whistle-blowers M2BS17P1
compromise tax plan F3ASD3P1
computer newspapers M1BS37P9
computer tax preparation M1BS07P1
computer trade group M2BS03P3
computer voting systems M2BS22P9
concord construction company M2BS14P6
consumer credit chain M1BS33P4
corner drug store F2BS11P2
credit scoring system M2BS30P1
dukakis budget chief M4BS01P3
emergency budget bill M4BS47P3
emergency evacuation plans M3BS06P5
emergency state aid F1AS03P4
exhibition baseball F1AS14P8
fox network F3AS41P1
government user fees M4BS60P6
harvard aids institute F2BS21P1
harvard law school F3AS90P1
holocaust remembrance week F3ASA9P1
home heating oil M4BS30P1
house appropriations committee M1BS32P3
hynes convention center M4BS61P1
iowa cornfield M3BS02P1
library operating costs F2BS33P1
livability stand point M4BS21P4
logan airport F3AS71P1
lotus law suit M1BS31P2
malden district court F3ASJ4P1
marine warrant officer F3ASS9P1
mass bankers association M2BS33P5
massachusetts amnesty period M4BS49P2
massachusetts ballot questions F3ASG5P1
massachusetts bar association F2BJRLP2
massachusetts biotechnology council M2BS26P1
massachusetts budget deficit F3AS47P1
massachusetts budget planners F1AS30P2
massachusetts businessmen M2BS35P5
massachusetts computer companies M2BS03P5
massachusetts computer industry M2BS03P9
massachusetts congressman F1AS35P2
massachusetts consumer groups F3ASF5P1
massachusetts correction system M1BS35P2
massachusetts corrections officials F1AS28P5
massachusetts credit worthiness F3AS75P1
massachusetts energy secretary F1AS39P2
massachusetts insurance division F3ASN2P1
massachusetts landlords M2BS24P1
massachusetts lawmaker F1AS31P2
massachusetts lieutenant governor F1AS37P2
massachusetts restaurant association M4BS04P3
massachusetts school children F1AS22P3
massachusetts state agencies F1AS36P2
massachusetts state house F3AS64P1
massachusetts wage earners M4BS58P1
massachusetts waste shipments M2BS05P1
metco classmates F2BS19P3
news headlines F2BS14P4
operation desert shield F3ASJ3PF
operation desert storm F3ASP7P1
police manhunt F2BS20P2

preseason celtics practice F3ASG3P1
raynham dog track M2BS23PA
rockingham horse track M2BS23PA
roxbury community activists M1BS21P1
roxbury community college F3AST9P1
roxbury housing project F2BS17P4
roxbury treatment center F3AS15P1
school drug use M1BS26P6
shrewsbury research center M2BS15P1
somerville mediation program M1BS19P2
state chairman M4BS24P1
state disaster relief F3ASV5P1
state education commissioner F1AS42P2
state energy secretary F1AS42P4
state excise taxes M2BS07P1
state health programs M4BS22P4
state income tax F3ASV6P1
state lawmakers F3AS77P1
state office complex M1BS20P2
state payroll F3AS44P1
state retirement administration M4BS13P3
state sales tax F3ASH6P1
state tax collections M4BS43P2
state taxpayers M2BS07P7
state transportation secretary M4BS59P3
superman comic book F3ASK6P1
tandy radio shack M2BS03P8
team locker room F3ASI9P1
telephone mouth piece M2BS25P7
tiffany network F3ASV3P1
town school system F3ASU8P1
traffic headaches F3AS21P1
trinity repertory company F3AS39P1
voice information service M1BS37P2
washington law professor F1AS34P2
welfare case worker M4BS07P6
wheaton classroom F3AS01P1
yale law school F3ASF4P1

Experiment 1

The table lists the 40 compounds constructed in order to test the predictions of the Embedded Prominence Hypothesis (EPH).

L/N1	<i>coffee table designer</i> <i>day care center</i> <i>field hockey player</i> <i>hay fever treatment</i> <i>kidney stone removal</i> <i>lung cancer surgery</i> <i>money market fund</i> <i>security guard service</i> <i>sign language class</i> <i>weather station data</i>	L/N2	<i>city hall restoration</i> <i>cotton candy maker</i> <i>cream cheese recipe</i> <i>diamond ring exhibition</i> <i>family planning clinic</i> <i>gene therapy technology</i> <i>maple syrup production</i> <i>science fiction book</i> <i>silicon chip manufacturer</i> <i>silver jubilee gift</i>
R/N2	<i>adult jogging suit</i> <i>business credit card</i> <i>celebrity golf tournament</i> <i>conference time sheet</i> <i>passenger test flight</i> <i>piano sheet music</i> <i>restaurant tourist guide</i> <i>student season ticket</i> <i>team locker room</i> <i>visitor name tag</i>	R/N3	<i>baby lemon tea</i> <i>company internet page</i> <i>family christmas dinner</i> <i>pilot leather jacket</i> <i>pizza home delivery</i> <i>prisoner community service</i> <i>student string orchestra</i> <i>tennis grass court</i> <i>tennis group practice</i> <i>woman fruit cocktail</i>

Test set presented to the participants of production experiment 1. The experiment tested the predictions of the Embedded Prominence Hypothesis (EPH). The set was presented to the participants in one of three different orders.

We are investigating the pronunciation of North American English and are making recordings of different speakers. Please fill in the following questionnaire before you start. Thank you very much for your cooperation.

List A

Subject #: _____ Age: _____

What are you majoring in? _____

Are you male or female? Male Female

Are you a native speaker of American English? Yes No

Would you consider yourself bilingual? Yes No

If yes, which other language besides English do you speak? _____

Since when have you lived in Toronto? _____

In which province(s) did you grow up? _____

Do you suffer from any speech or hearing disorders? Yes No

Instructions

Please read out the following sentences and words as naturally as possible. We will record your voice. If you stutter, hesitate or make a mistake, please read out the whole sentence again. The reading will take about 8 to 10 minutes. If you need a break during your reading, please give a sign to the researcher so that she can pause the recording.

Before the actual recording starts, please read out the following five test sentences. While you are reading, the researcher will adjust the recorder to your voice. You may have to read these sentences twice. The test sentences are also there for you to get used to the test situation as well as to the action of reading out loud.

Test sentences

1) Sandra is visiting her best friend in Paris. 2) Her friend moved to Paris two years ago. 3) This is the first time that Sandra is flying to Europe. 4) She really looks forward to it. 5) Her grandparents told her all about Paris last week.

If you have any questions, please ask them now. The actual experiment starts on the next page. When you are ready, tell the researcher to start the recording!

He watched an old western movie last night.

He rented a nice beach apartment last summer.

She started hay fever treatment last year.

He organized an exciting adventure vacation last week.

He wrote a restaurant tourist guide last month.

He booked a non-stop transatlantic flight last week.

He sold a cotton candy maker last month.

She bought a small wooden horse for her daughter last week.

She voted for a city hall restoration last month.

She attended a Spanish and French class last semester.

He had a lung cancer surgery last year.

She played a Japanese card game with her family last night.

He ordered an adult jogging suit last week.

She made a beautiful farewell gift for her friend last night.

She founded a student string orchestra last month.

She looked at a black and white photo of her grandmother last night.

She read about a coffee table designer last week.

She saw a little green bird last Friday.

He signed up for a business credit card last month.

She was a good and ambitious student last year.

He bought a science fiction book last Friday.

She ordered a non-alcoholic cocktail last night.

He started tennis group practice last month.

She married a big hairy guy last weekend.

She attended a sign language class last week.

He interviewed a track and field athlete last Tuesday.

She worked at a day care centre last year.

She visited her Mexican-American family last June.

She participated in a passenger test flight last year.

She played hide and seek with her children last weekend.

He participated in a celebrity golf tournament last year.

He lost his hand-made scarf last Monday.

He spoke to a silicon chip manufacturer last month.

She learned a second language last year.

He missed the family Christmas dinner last night.

He cancelled an important business appointment last month.

He tasted some baby lemon tea last week.

He met a crazy pop artist last Monday.

He received some weather station data last night.

She listened to a new country band last evening.

He had a kidney stone removal last Friday.

She had good and bad times last year.

She designed a team locker room last week.

She painted a big green apple at school last Tuesday.

He bought a silver jubilee gift last week.

He sang a wonderful love song last night.

She designed a pilot leather jacket last month.

He wrote a sad and emotional short story last month.

She visited a diamond ring exhibition last July.

He called his best female friend last night.

He practiced for the national long-jump championships last March.

She lost her visitor name tag last night.

He cooked delicious French food last evening.

He invested in a money market fund last month.

He called an old school friend last weekend.

She married a field hockey player last Sunday.

She read a funny comic book last night.

She bought a student season ticket last week.

He sold an expensive sailing boat last month.

She made a conference time sheet last Monday.

She applied for a well-paid job at a bank last month.

He worked at a family planning clinic last year.

She travelled to a small tropical island last August.

He read about maple syrup production last week.

He learned some difficult new words in Swedish last night.

She was sentenced to prisoner community service last year.

She missed her favourite TV show last night.

He played on a tennis grass court last month.

He met new school friends last week.

She read about a gene therapy technology last night.

He cooked a spicy Indian meal last Monday.

He looked at a company internet page last Tuesday.

He drove a slow yellow bus last year.

She searched for a cream cheese recipe last week.

She met a nice French guy last month.

She called a pizza home delivery last night.

He lost his black leather wallet last month.

She lost her golden necklace last month.

He looked for piano sheet music last Monday.

She slept on a comfortable blue coach last night.

She hired a security guard service last month.

She ordered a woman fruit cocktail last night.

Thank you very much!

Experiment 2

The four tables list the 42 compounds constructed in order to test the predictions of the Immediate Constituent Prominence Hypothesis (IPH).

avenue compounds	street compounds
<i>[Bird Rock] Avenue</i>	<i>[Bird Rock] Street</i>
<i>[Post office] Avenue</i>	<i>[Post office] Street</i>
<i>[Lemon Grove] Avenue</i>	<i>[Lemon Grove] Street</i>
<i>[Ocean Front] Avenue</i>	<i>[Ocean Front] Street</i>
<i>[Fish market] Avenue</i>	<i>[Fish market] Street</i>

pie compounds	argument-head compounds
<i>[passion fruit] pie</i>	<i>[passion fruit] seller</i>
<i>[candy bar] pie</i>	<i>[candy bar] buyer</i>
<i>[blood orange] pie</i>	<i>[blood orange] supplier</i>
<i>[orange juice] pie</i>	<i>[orange juice] producer</i>
<i>[potato chip] pie</i>	<i>[potato chip] taster</i>

IC2 during IC1	argument-head compounds
<i>[field trip] weather</i>	<i>[field trip] leader</i>
<i>[winter term] breakfast</i>	<i>[winter term] preparation</i>
<i>[Monday morning] meeting</i>	<i>[Monday morning] hater</i>
<i>[flight test] accident</i>	<i>[flight test] manual</i>
<i>[spring break] vacation</i>	<i>[spring break] organization</i>
<i>[holiday season] job</i>	<i>[holiday season] planner</i>

IC2 is located at IC1	argument-head compounds
<i>[cocktail bar] fight</i>	<i>[cocktail bar] designer</i>
<i>[coffee house] concert</i>	<i>[coffee house] lover</i>
<i>[gas station] robbery</i>	<i>[gas station] operator</i>
<i>[apartment building] party</i>	<i>[apartment building] manager</i>
<i>[hockey stadium] event</i>	<i>[hockey stadium] cleaner</i>
<i>[movie theater] fire</i>	<i>[movie theater] visitor</i>

Test set presented to the participants of production experiment 2. The experiment tested the predictions of the Immediate Constituent Prominence Hypothesis (IPH). The set was presented to the participants in one of three different orders.

We are investigating the pronunciation of North American English and are making recordings of different speakers. Please fill in the following questionnaire before you start. Thank you very much for your cooperation.

List A

Subject #: _____ Age: _____

What are you majoring in? _____

Are you male or female? Male Female

Are you a native speaker of American English? Yes No

Would you consider yourself bilingual? Yes No

If yes, which other language besides English do you speak? _____

Since when have you lived in Toronto? _____

In which province(s) did you grow up? _____

Do you suffer from any speech or hearing disorders? Yes No

Instructions

Please read out the following sentences and words as naturally as possible. We will record your voice. If you stutter, hesitate or make a mistake, please read out the whole sentence again. The reading will take about 8 to 10 minutes. If you need a break during your reading, please give a sign to the researcher so that she can pause the recording.

Before the actual recording starts, please read out the following five test sentences. While you are reading, the researcher will adjust the recorder to your voice. You may have to read these sentences twice. The test sentences are also there for you to get used to the test situation as well as to the action of reading out loud.

Test sentences

1) Sandra is visiting her best friend in Paris. 2) Her friend moved to Paris two years ago. 3) This is the first time that Sandra is flying to Europe. 4) She really looks forward to it. 5) Her grandparents told her all about Paris last week.

If you have any questions, please ask them now. The actual experiment starts on the next page. When you are ready, tell the researcher to start the recording!

He made an orange juice pie last week.

She organized a great pool party last year.

She lived at Bird Rock Avenue last semester.

She released her latest rock album last month.

He heard of a cocktail bar fight last night.

He bought a brand new car last Monday.

He looked for Ocean Front Street last Friday.

She worked as a part-time waitress last year.

She organized a hockey stadium event last month.

He bought a one-way ticket to Sydney last week.

He applied for a holiday season job last summer.

She fell in love with a good-looking doctor last year.

He enjoyed the field trip weather last Tuesday.

He downloaded some soul and rap music last night.

She was called a Monday morning hater last week.

She worried a lot about her future professional career last night.

He lived at Post office Street last year.

She lived in a big cheap apartment last summer.

He became a potato chip taster last Monday.

He met a do-it-yourself enthusiast last Saturday.

She went to Fish market Avenue last Tuesday.

She threw a real big dinner party two weeks ago.

She met an orange juice producer last week.

She read a short but very good article last night.

He lived at Bird Rock Street last semester.

She was a good and loving mother in the past.

He read about a flight test accident last night.

She learned some difficult new dance moves last week.

He ate a candy bar pie last night.

He bought a nice glass bowl last Monday.

He looked for Ocean Front Avenue last Friday.

She went to a nearby post office last week.

He worked as a cocktail bar designer last year.

She participated in an anti-war protest last June.

He witnessed a gas station robbery last night.

She wore a short red dress last evening.

She went to an apartment building party last week.

She went to a new cocktail bar last night.

He spoke to a field trip leader last Monday.

She participated in a Monday morning meeting last week.

He bought a small gas station last month.

He worked for a spring break organization last year.

He forgot her sweet-sixteen birthday party last week.

He drove into the wrong Street direction this morning.

She lived at Post office Avenue last year.

He made his first strawberry cake last Friday.

He met a candy bar buyer last Wednesday.

He went to a lovely Irish pub last Saturday.

He went to a coffee house concert last month.

He walked from University to Park Street yesterday.

He asked for Lemon Grove Street last night.

He washed his dirty sport clothes last weekend.

She spoke to a movie theatre visitor last night.

He moved into a small campus apartment four days ago.

He was an apartment building manager last month.

She met her first boyfriend last week.

He worked as a holiday season planner last summer.

She wore a long green skirt last week.

He went to Fish market Street last Tuesday.

She organized a lovely surprise party for her sister last weekend.

He tried a potato chip pie last Monday.

He had a full time job last year.

He spoke to a passion fruit seller last Wednesday.

He moved into an old smelly apartment last week.

He met a gas station operator last night.

He invented an automatic potato peeler two years ago.

He signed up for a winter term breakfast last week.

He swam at an inner-city school tournament two years ago.

He hired a hockey stadium cleaner last month.

He visited his former history teacher last year.

He ate a passion fruit pie last Wednesday.

He watched "Good morning Vietnam" this morning.

She talked to a blood orange supplier last week.

She travelled with her best friend last November.

He was a coffee house lover last year.

He became a successful and rich business man four years ago.

She enjoyed her spring break vacation last April.

He went to a new fish market yesterday.

He went to Lemon Grove Avenue last night.

She spoke to a young and handsome teacher last evening.

She read about a movie theatre fire last week.

She made a long-distance phone call last night.

He studied a flight test manual last Monday.

She bought fresh vegetables yesterday.

He started his winter term preparation last month.

He found a really good cake recipe last night.

She made a blood orange pie last week.

Thank you very much

B Mixed-effects models, regression models and likelihood ratio tests

Mixed-effects models

Corpus study

Final mixed-effects model of corpus study showing the full random effects structure and the fixed effects structure of the model. Baseline: lpitch, left, female. The model's baseline was changed by means of treatment contrasts.

```
Linear mixed model fit by REML
Formula: pitchST ~ Branch * position + gender + (1|speaker) + (1|item)
Data: BURSCout
   AIC   BIC logLik deviance REMLdev
6527 6579 -3254   6506   6507
Random effects:
Groups   Name      Variance Std.Dev.
item    (Intercept) 2.3796   1.5426
speaker (Intercept) 3.5578   1.8862
Residual                    7.0743   2.6598
Number of obs: 1290, groups: item, 447; speaker, 7

Fixed effects:
              Estimate Std. Error t-value p-value
(Intercept)  20.62265    1.10940  18.589 < 0.001
right        -0.02123    0.33060  -0.064 < 0.949
mpitch       -2.00775    0.21217  -9.463 < 0.001
rpitch       -2.68423    0.21491 -12.490 < 0.001
male         -6.39112    1.45958  -4.379 < 0.001
right:mpitch  1.64564    0.40340   4.079 < 0.001
right:rpitch -1.05698    0.41019  -2.577 < 0.05
```

Final mixed-effects model of corpus study showing the full random effects structure and the fixed effects structure of the model. Baseline: rpitch, left, female. The model's baseline was changed by means of treatment contrasts.

```
Linear mixed model fit by REML
Formula: pitchST ~ Branch * position + gender + (1|speaker) + (1|item)
Data: BURSCout
   AIC   BIC logLik deviance REMLdev
6527 6579 -3254   6506   6507
Random effects:
Groups   Name      Variance Std.Dev.
item    (Intercept) 2.3796   1.5426
speaker (Intercept) 3.5578   1.8862
```

Residual 7.0743 2.6598
 Number of obs: 1290, groups: item, 447; speaker, 7

Fixed effects:

	Estimate	Std. Error	t-value	p-value
(Intercept)	17.9384	1.1099	16.162	< 0.001
right	-1.0782	0.3381	-3.189	< 0.01
lpitch	2.6842	0.2149	12.490	< 0.001
mpitch	0.6765	0.2148	3.149	< 0.01
male	-6.3911	1.4596	-4.379	< 0.001
right:lpitch	1.0570	0.4102	2.577	< 0.5
right:mpitch	2.7026	0.4094	6.601	< 0.001

Experiment 1

Final mixed-effects model of experiment 1 showing the full random effects structure and the fixed effects structure of the model. Baseline: mpitch, lrn2, female. The model's baseline was changed by means of treatment contrasts.

Linear mixed model fit by REML
 Formula: pitchST ~ Branching * position + position:gender + gender + (1|speaker) + (1|item)
 Data: triabsmfout
 AIC BIC logLik deviance REMLdev
 3715 3793 -1843 3666 3685
 Random effects:
 Groups Name Variance Std.Dev.
 item (Intercept) 0.20633 0.45424
 speaker (Intercept) 9.70449 3.11520
 Residual 0.85526 0.92481
 Number of obs: 1308, groups: item, 40; speaker, 12

Fixed effects:

	Estimate	Std. Error	t-value	p-value
(Intercept)	30.1735	1.1838	25.490	< 0.001
ln1	-0.6681	0.2056	-3.250	< 0.01
rn3	-0.1449	0.2063	-0.703	0.483
lpitch	0.7291	0.1016	7.175	< 0.001
rpitch	-1.8116	0.1054	-17.189	< 0.001
male	-10.3700	1.8259	-5.679	< 0.001
ln1:lpitch	0.8341	0.1514	5.511	< 0.001
rn3:lpitch	0.5843	0.1514	3.860	< 0.001
ln1:rpitch	0.6220	0.1541	4.036	< 0.001
rn3:rpitch	0.5075	0.1549	3.276	< 0.01
lpitch:male	0.1124	0.1257	0.895	0.373
rpitch:male	0.3724	0.1280	2.909	< 0.01

Regression model of excluded speaker. Baseline: mpitch, ln1, female. The model's baseline was changed by means of treatment contrasts.

Residuals:
 Min 1Q Median 3Q Max
 -4.9807 -1.2836 -0.1980 1.6568 6.4478

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	15.6126	0.7172	21.768	< 2e-16 ***
ln2	6.5491	0.9886	6.624	1.83e-09 ***
rn2	5.6663	0.9886	5.731	1.08e-07 ***
rn3	6.2677	0.9886	6.340	6.88e-09 ***
lpitch	6.7307	1.0143	6.636	1.74e-09 ***


```

rpitch          0.4353      1.0456   0.416 0.678083
ln2:lpitch     -6.8796      1.3982  -4.920 3.44e-06 ***
rn2:lpitch     -6.2427      1.3982  -4.465 2.13e-05 ***
rn3:lpitch     -5.8366      1.4164  -4.121 7.85e-05 ***
ln2:rpitch     -6.0694      1.4210  -4.271 4.47e-05 ***
rn2:rpitch     -5.2261      1.4892  -3.509 0.000677 ***
rn3:rpitch     -1.5277      1.4390  -1.062 0.290981
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.152 on 99 degrees of freedom
Multiple R-squared:  0.6351,    Adjusted R-squared:  0.5946
F-statistic: 15.67 on 11 and 99 DF,  p-value: < 2.2e-16

```

Regression model of excluded speaker. Baseline: rpitch, rn2. The model's baseline was changed by means of treatment contrasts.

```

Call:
aov(formula = pitchST ~ branching * position, data = triabs25out)

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-4.9807 -1.2836 -0.1980  1.6568  6.4478

```

```

Coefficients:
            Estimate Std. Error t-value Pr(>|t|)
(Intercept)  16.48812   0.81328  20.274 < 2e-16 ***
ln1          -0.44020   1.11362  -0.395 0.693482
ln2           0.03955   1.06038   0.037 0.970322
rn3           4.29987   1.08437   3.965 0.000139 ***
lpitch        5.27885   1.06038   4.978 2.71e-06 ***
mpitch        4.79085   1.06038   4.518 1.73e-05 ***
ln1:lpitch    1.01661   1.48916   0.683 0.496405
ln2:lpitch    0.20639   1.43192   0.144 0.885685
rn3:lpitch   -3.29232   1.46741  -2.244 0.027084 *
ln1:mpitch   -5.22613   1.48916  -3.509 0.000677 ***
ln2:mpitch    0.84323   1.43192   0.589 0.557282
rn3:mpitch   -3.69847   1.44977  -2.551 0.012270 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 2.152 on 99 degrees of freedom
Multiple R-squared:  0.6351,    Adjusted R-squared:  0.5946
F-statistic: 15.67 on 11 and 99 DF,  p-value: < 2.2e-16

```

Experiment 2

Mixed-effects model for subset [*winter term* + N3]. Baseline: lpitch, duration, female. The model's baseline was changed by means of treatment contrasts.

```

Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: winterout
    AIC   BIC logLik deviance REMLdev
 246.6 268.6 -114.3   224.2   228.6
Random effects:
 Groups   Name      Variance Std.Dev.
 speaker (Intercept) 4.86864  2.20650
 Residual              0.39629  0.62952
Number of obs: 86, groups: speaker, 16

```

```

Fixed effects:
              Estimate Std. Error t-value p-value
(Intercept)  22.67285   0.75540  30.015 < 0.001
mpitch       0.04614   0.23493   0.196  0.816
rpitch      -1.19162   0.24060  -4.953 < 0.001
control      1.16599   0.22744   5.127 < 0.001
male        -8.70381   1.12099  -7.764 < 0.001
mpitch:control -0.85860   0.32677  -2.628 < 0.05
rpitch:control -1.03751   0.33724  -3.077 < 0.05

```

Mixed-effects model for subset [*holiday season* + N3]. Baseline: lpitch, location, female.

```

Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: holidayout
      AIC      BIC logLik deviance REMLdev
261.8 284.3 -121.9  239.8  243.8
Random effects:
Groups  Name      Variance Std.Dev.
speaker (Intercept) 5.48489  2.34199
Residual              0.42567  0.65243
Number of obs: 90, groups: speaker, 16

```

```

Fixed effects:
              Estimate Std. Error t-value p-value
(Intercept)  22.55503   0.79986  28.199 < 0.001
mpitch      -0.00562   0.23523  -0.024  0.981
rpitch      -1.95116   0.24101  -8.096 < 0.001
control      0.89669   0.23067   3.887 < 0.001
male        -8.74372   1.18856  -7.357 < 0.001
mpitch:control -0.51390   0.33315  -1.543  0.127
rpitch:control  0.21063   0.33767   0.624  0.535

```

Mixed-effects model for subset [*holiday season* + N3]. Baseline: rpitch, location, female.

The model's baseline was changed by means of treatment contrasts.

```

Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: holidayout
      AIC      BIC logLik deviance REMLdev
261.8 284.3 -121.9  239.8  243.8
Random effects:
Groups  Name      Variance Std.Dev.
speaker (Intercept) 5.48489  2.34199
Residual              0.42567  0.65243
Number of obs: 90, groups: speaker, 16

```

```

Fixed effects:
              Estimate Std. Error t-value p-value
(Intercept)  20.6039   0.8029  25.662 < 0.001
lpitch       1.9512   0.2410   8.096 < 0.001
mpitch       1.9455   0.2457   7.917 < 0.001
control      1.1073   0.2466   4.490 < 0.001
male        -8.7437   1.1886  -7.357 < 0.001
lpitch:control -0.2106   0.3377  -0.624  0.535
mpitch:control -0.7245   0.3444  -2.104  0.038

```

Mixed-effects model for subset [*spring break* + N3]. Baseline: rpitch, location, female.

The model's baseline was changed by means of treatment contrasts.

```

Linear mixed model fit by REML

```

```

Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: springout
AIC BIC logLik deviance REMLdev
276.6 298.3 -129.3 257.5 258.6
Random effects:
Groups Name Variance Std.Dev.
speaker (Intercept) 5.99622 2.44872
Residual 0.68202 0.82584
Number of obs: 83, groups: speaker, 16

```

```

Fixed effects:
Estimate Std. Error t-value p-value
(Intercept) 22.0653 0.8539 25.839 < 0.001
lpitch 2.0137 0.3239 6.217 < 0.001
mpitch 0.4214 0.3239 1.301 0.197
control -0.6476 0.3546 -1.826 0.072
male -8.9933 1.2486 -7.203 < 0.001
lpitch:control 1.0683 0.4617 2.314 0.024
mpitch:control 1.2116 0.4617 2.625 < 0.05

```

Mixed-effects model for subset [*flight test* + N3]. Baseline: mpitch, location, female.
The model's baseline was changed by means of treatment contrasts.

```

Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: flightout
AIC BIC logLik deviance REMLdev
309.5 331.6 -145.8 292.1 291.5
Random effects:
Groups Name Variance Std.Dev.
speaker (Intercept) 5.7908 2.4064
Residual 1.0029 1.0014
Number of obs: 86, groups: speaker, 16

```

```

Fixed effects:
Estimate Std. Error t-value p-value
(Intercept) 22.3898 0.8526 26.261 < 0.001
lpitch 0.7858 0.3871 2.030 < 0.05
rpitch -1.4505 0.3969 -3.655 < 0.001
control 1.3667 0.3703 3.691 < 0.001
male -8.4485 1.2332 -6.851 < 0.001
lpitch:control -1.0893 0.5246 -2.076 < 0.05
rpitch:control -0.9167 0.5369 -1.707 < 0.092

```

Mixed-effects model for subset [*coffee house* + N3]. Baseline: lpitch, location, female.
The model's baseline was changed by means of treatment contrasts.

```

Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: coffeeout
AIC BIC logLik deviance REMLdev
338.7 361.1 -160.3 322.2 320.7
Random effects:
Groups Name Variance Std.Dev.
speaker (Intercept) 4.8528 2.2029
Residual 1.3523 1.1629
Number of obs: 89, groups: speaker, 17

```

```

Fixed effects:
Estimate Std. Error t-value p-value
(Intercept) 24.8171 0.7615 32.59 < 0.001

```

B Mixed-effects models, regression models and likelihood ratio tests

mpitch	-2.2974	0.4303	-5.34 < 0.001
rpitch	-2.7369	0.4111	-6.66 < 0.001
control	-0.9980	0.4255	-2.35 < 0.05
male	-8.9756	1.1173	-8.03 < 0.001
mpitch: control	1.2649	0.6110	2.07 < 0.05
rpitch: control	0.1948	0.5978	0.33 0.745

Mixed-effects model for subset [*apartment building* + N3]. Baseline: rpitch, location, female. The model's baseline was changed by means of treatment contrasts.

```
Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: apartmentout
   AIC   BIC logLik deviance REMLdev
336.1 358.8 -159.1   319.1   318.1
Random effects:
 Groups   Name      Variance Std.Dev.
speaker (Intercept) 8.73944  2.95625
Residual                    0.99846  0.99923
Number of obs: 92, groups: speaker, 17

Fixed effects:
              Estimate Std. Error t-value p-value
(Intercept)    21.2905    0.9789  21.749 < 0.001
lpitch          1.6980    0.3777   4.496 < 0.001
mpitch          0.3029    0.3777   0.802  0.425
control        -0.8348    0.3662  -2.280  0.025
male           -8.2297    1.4730  -5.587 < 0.001
lpitch: control 1.4585    0.5143   2.836 < 0.01
mpitch: control 1.4420    0.5100   2.827 < 0.01
```

Mixed-effects model for subset [*hockey stadium* + N3]. Baseline: rpitch, location, female. The model's baseline was changed by means of treatment contrasts.

```
Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: hockeyout
   AIC   BIC logLik deviance REMLdev
365.2 388.4 -173.6   348.3   347.2
Random effects:
 Groups   Name      Variance Std.Dev.
speaker (Intercept) 5.5995  2.3663
Residual                    1.2969  1.1388
Number of obs: 97, groups: speaker, 17

Fixed effects:
              Estimate Std. Error t-value p-value
(Intercept)    19.7816    0.8126  24.343 < 0.001
lpitch          2.6576    0.4131   6.433 < 0.001
mpitch          2.1146    0.4212   5.020 < 0.001
control          1.6706    0.4058   4.116 < 0.001
male           -8.4130    1.1894  -7.074 < 0.001
lpitch: control -1.0002    0.5685  -1.759  0.082
mpitch: control -2.2173    0.5745  -3.860 < 0.001
```

Mixed-effects model for subset [*hockey stadium* + N3]. Baseline: lpitch, location, female. The model's baseline was changed by means of treatment contrasts.

```

Linear mixed model fit by REML
Formula: pitchST ~ position * semantics + gender + (1|speaker)
Data: hockeyout
AIC   BIC logLik deviance REMLdev
365.2 388.4 -173.6   348.3   347.2
Random effects:
Groups   Name      Variance Std.Dev.
speaker (Intercept) 5.5995   2.3663
Residual                1.2969   1.1388
Number of obs: 97, groups: speaker, 17

Fixed effects:
                Estimate Std. Error t-value p-value
(Intercept)      21.8962    0.8126  26.945 < 0.001
lpitch            0.5430    0.4131   1.314  0.192
rpitch           -2.1146    0.4212  -5.020 < 0.001
control          -0.5468    0.4058  -1.347  0.181
male             -8.4130    1.1894  -7.074 < 0.001
lpitch:control    1.2171    0.5685   2.141  0.035
rpitch:control    2.2173    0.5745   3.860 < 0.001

```

Likelihood ratio test

Corpus study

Likelihood ratio test for fully specified model (BURSC.lmer) and reduced model (BURSC-a) without random effect ITEM.

```

Models:
BURSC-a.lmer: pitchST ~ Branch * position + gender + (1|speaker)
BURSC.lmer: pitchST ~ Branch * position + gender + (1|speaker) + (1|item)
                Df  AIC    BIC   logLik Chisq Chi Df Pr(>Chisq)
BURSC-a.lmer    9 6793.6 6840.2 -3387.8
BURSC.lmer     10 6749.5 6801.3 -3364.8 46.075    1 1.138e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Likelihood ratio test for fully specified model (BURSC.lmer) and reduced model (BURSC-b) without random effect SPEAKER.

```

BURSC-b.lmer: pitchST ~ Branch * position + gender + (1|item)
BURSC.lmer: pitchST ~ Branch * position + gender + (1|speaker) + (1|item)
                Df  AIC    BIC   logLik Chisq Chi Df Pr(>Chisq)
BURSC-b.lmer3   9 6919.9 6966.5 -3451.0
BURSC.lmer     10 6749.5 6801.3 -3364.8 172.4    1 < 2.2e-16 ***

```

Experiment 1

Likelihood ratio test for fully specified model (EPH.lmer) and reduced model (EPH-a.lmer) without random effect SPEAKER.

```

EPH-a.lmer: pitchST ~ branching * position * gender + (1|item)
EPH.lmer: pitchST ~ branching * position * gender + (1|speaker) + (1|item)
                Df  AIC    BIC   logLik Chisq Chi Df Pr(>Chisq)
EPH-a.lmer     26 6593.6 6728.2 -3270.8
EPH.lmer       27 3712.8 3852.6 -1829.4 2882.8    1 < 2.2e-16 ***

```

Likelihood ratio test for fully specified model (EPH.lmer) and reduced model (EPH-b.lmer) without random effect ITEM.

```
EPH-b.lmer: pitchST ~ branching * position * gender + (1|speaker)
EPH.lmer: pitchST ~ branching * position * gender + (1|speaker) + (1|item)
      Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
EPH-b.lmer  26 3884.3 4018.9 -1916.2
EPH.lmer    27 3712.8 3852.6 -1829.4 173.47      1 < 2.2e-16 ***
```

Likelihood ratio test for model in which ln1 and rn2 are treated as two separate factor levels (LN2RN2-sep) and reduced model in which the levels ln2 and rn2 are combined to form the single factor level lrn2 (LN2RN2-one).

```
ln2rn2-one.lmer: pitchST ~ branching * position + position:gender + gender + (1|speaker) + (1|item)
ln2rn2-sep.lmer: pitchST ~ branching * position + position:gender + gender + (1|speaker) + (1|item)
      Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
ln2rn2-one.lmer  15 3695.9 3773.6 -1833.0
ln2rn2-sep.lmer  18 3697.8 3791.0 -1830.9 4.1031      3 0.2505
```

Experiment 2

Likelihood ratio test for fully specified model (avenue.lmer) and reduced model (avenue-a.lmer) without random effect COMPLEX.

```
avenue-a.lmer.lmer: pitchST ~ position * analogy * gender + (1|speaker)
avenue.lmer.lmer: pitchST ~ position * analogy * gender + (1|speaker) + (1|complex)
      Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
avenue-a.lmer  14 1675.4 1733.4 -823.69
avenue.lmer    15 1626.5 1688.6 -798.25 50.892      1 9.76e-13 ***
```

Likelihood ratio test for fully specified model (avenue.lmer) and reduced model (avenue-b.lmer) without random effect SPEAKER.

```
avenue-b.lmer: pitchST ~ position * analogy * gender + (1|complex)
avenue.lmer: pitchST ~ position * analogy * gender + (1|speaker) + (1|complex)
      Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
avenue-b.lmer  14 2197.5 2255.5 -1084.76
avenue.lmer    15 1626.5 1688.6 -798.25 573.02      1 < 2.2e-16 ***
```

Likelihood ratio test for fully specified model (pie.lmer) and reduced model (pie-b.lmer) without random effect COMPLEX.

```
pie-a.lmer: pitchST ~ position * analogy * gender + (1|speaker)
pie.lmer: pitchST ~ position * analogy * gender + (1|speaker) + (1|complex)
      Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
pie-a.lmer  14 1816.9 1874.5 -894.44
pie.lmer    15 1811.6 1873.3 -890.80 7.2929      1 0.006923 **
```

Likelihood ratio test for fully specified model (pie.lmer) and reduced model (pie-b.lmer) without random effect SPEAKER.

```

pie-b.lmer: pitchST ~ position * analogy * gender + (1|complex)
pie.lmer: pitchST ~ position * analogy * gender + (1|speaker) + (1|complex)
           Df   AIC   BIC logLik  Chisq Chi Df Pr(>Chisq)
pie-b.lmer 14 2253.5 2311.1 -1112.8
pie.lmer   15 1811.6 1873.3  -890.8 443.92    1 < 2.2e-16 ***

```

Likelihood ratio test for fully specified model (during.lmer) and reduced model (during -a.lmer) without random effect COMPLEX.

```

duration-a.lmer: pitchST ~ position * semantics * gender + (1|speaker)
duration.lmer: pitchST ~ position * semantics * gender + (1|speaker) + (1|complex)
           Df   AIC   BIC logLik  Chisq Chi Df Pr(>Chisq)
duration-a.lmer      14 1950.8 2010.8 -961.37
duration.lmer        15 1938.4 2002.7 -954.22 14.309    1 0.0001551 ***

```

Likelihood ratio test for fully specified model (during.lmer) and reduced model (during-b.lmer) without random effect SPEAKER.

```

duration-b.lmer: pitchST ~ position * semantics * gender + (1|complex)
duration.lmer: pitchST ~ position * semantics * gender + (1|speaker) + (1|complex)
           Df   AIC   BIC logLik  Chisq Chi Df Pr(>Chisq)
duration-b.lmer      14 2547.6 2607.6 -1259.78
duration.lmer        15 1938.4 2002.7 -954.22 611.12    1 < 2.2e-16 ***

```

Likelihood ratio test for fully specified model (location.lmer) and reduced model (location-a.lmer) without random effect COMPLEX.

```

location-a.lmer: pitchST ~ position * semantics * gender + (1|speaker)
location.lmer: pitchST ~ position * semantics * gender + (1|speaker) + (1|complex)
           Df   AIC   BIC logLik  Chisq Chi Df Pr(>Chisq)
location-a.lmer      14 2139.3 2200.1 -1055.6
location.lmer        15 2129.5 2194.7 -1049.8 11.747    1 0.0006093 ***

```

Likelihood ratio test for fully specified model (location.lmer) and reduced model (location-b.lmer) without random effect SPEAKER.

```

location-b.lmer: pitchST ~ position * semantics * gender + (1|complex)
location.lmer: pitchST ~ position * semantics * gender + (1|speaker) + (1|complex)
           Df   AIC   BIC logLik  Chisq Chi Df Pr(>Chisq)
location-b.lmer      14 2813.7 2874.5 -1392.9
location.lmer        15 2129.5 2194.7 -1049.8 686.2    1 < 2.2e-16 ***

```


Bibliography

- Baayen, H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge University Press.
- Baayen, R. H. and Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3.2:12–28.
- Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX lexical database*. Linguistic Data Consortium, Philadelphia, PA.
- Bates, D. M. and Sarkar, D. (2007). lme4: Linear mixed-effects models using S4 classes (version 0.9975-13). R package. Retrieved 27 March 2007, from <http://lib.stat.cmu.edu/R/CRAN/src/contrib/Descriptions/lme4.html>.
- Bauer, L. (1983a). *English word-formation*. Cambridge University Press, Cambridge.
- Bauer, L. (1983b). Stress in compounds: a rejoinder. *English Studies*, 64:47–53.
- Bauer, L. (1998). When is a sequence of two nouns a compound in English? *English Language and Linguistics*, 2(1):65–86.
- Bell, M. (2008). Noun noun constructions and the assignment of stress. Paper presented at the 1st conference of the international society for the linguistics of English (ISLE 1).
- Bell, M. (2011). At the boundary of morphology and syntax: noun noun constructions in English. In Galani, A., Hicks, G., and Tsoulus, G. (eds.), *Morphology and its Interfaces*, pp. 137–168. John Benjamins, Amsterdam and Philadelphia.
- Bell, M. and Plag, I. (2012). Informativeness is a determinant of compound stress in English. *Journal of Linguistics*. Available on CJO 2012 doi:10.1017/S0022226712000199.
- Berg, T. (2006). The internal structure of four-noun compounds in English and German. *Corpus Linguistics and Linguistic Theory*, 2(2):197–231.

- Berg, T. (2009). *Structure in language. A dynamic perspective*. Routledge, New York.
- Bloomfield, L. (1933). *Language*. Holt, New York.
- Boersma, P. and Weenik, D. J. (2007). Praat: doing phonetics by computer (version 4.6.36) [computer program]. Technical report, Amsterdam, Available at: <http://www.praat.org/>.
- Bullon, S. (ed.) (2002). *Longman dictionary of American English. Your complete Guide to American English*. Longman, White Plains, NY.
- Carstairs-McCarthy, A. (2002). *An introduction to English Morphology*. Edinburgh University Press, Edinburgh.
- Carvajal, C. S. and Horwood, J. (1996). *The Oxford Spanish-English dictionary: New international edition*. Oxford University Press, Oxford.
- Chomsky, N. and Halle, M. (1968). *The sound pattern of English*. Harper and Row, New York.
- Cinque, G. (1993). A null theory of phrase and compound stress. *Linguistic Inquiry*, 24(2):239–297.
- Collier, R. (1975). Physiological correlates of intonation patterns. *Journal of the Acoustical Society of America*, 58:249–255.
- Crawley, M. J. (2005a). *Statistical Computing: An Introduction to Data Analysis using S-Plus*. Wiley, 4 edition.
- Crawley, M. J. (2005b). *Statistics: An introduction to using R*. Wiley.
- Daelemans, W., Zavrel, J., van der Sloot, K., and van den Bosch, A. (2004). TiMBL: Tilburg Memory Based Learner (version 5.1). Computer program. Reference Guide ILK Technical Report 04 02, available from <http://ilk.uvt.nl/downloads/pub/papers/ilk0402.pdf>.
- Farnetani, E., Torsello, C. T., and Cosi, P. (1988). English compound versus non-compound noun phrases in discourse: An acoustic and perceptual study. *Language and Speech*, 31:157–180.
- Fudge, E. (1984). *English word stress*. George Allen and Unwin, London.

- Giegerich, H. J. (1985). *Metrical phonology and phonological structure*. Cambridge University Press, Cambridge.
- Giegerich, H. J. (1992). *English Phonology: An introduction*. Cambridge University Press, Cambridge.
- Giegerich, H. J. (2004). Compound or phrase? English noun-plus-noun constructions and the stress criterion. *English Language and Linguistics*, 8:1–24.
- Giegerich, H. J. (2009). The English compound stress myth. *Word formation*, 2(1):1–17.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press, Cambridge.
- Gut, U. (2009). *Introduction to English phonetics and phonology*. Lang, Frankfurt am Main.
- Hayes, B. (1995). *Metrical stress theory. Principles and case studies*. Chicago University of Chicago Press, Chicago.
- Henton, C. G. (1989). Fact and fiction in the description of female and male pitch. *Language and Communication*, 9:299–311.
- Hirschberg, J. (2002). Communication and prosody: functional aspects of prosody. *Speech Communication*, 36(1):31–43.
- Hoeksema, J. (1985). *Categorical morphology*. Garland Publisher, New York.
- Hornby, A. S. (1983). *Oxford's student's dictionary of American English*. Oxford University Press, Oxford.
- Hornby, A. S. (1995). *Oxford Advanced Learner's Dictionary New Edition*. Oxford University Press, Oxford.
- House, A. S. and Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, 25(1):103–113.
- Ingram, J., Nguyen, T. A. T., and Pensalfini, R. (2003). An acoustic analysis of compound and phrasal stress patterns in Australian English. Unpublished.
- Kingdon, R. (1958). *The groundwork of English stress*. Longmans, Green and Co., London.

- Kösling, K. and Plag, I. (2009). Does branching direction determine prominence assignment? An empirical investigation of triconstituent compounds in English. *Corpus Linguistics and Linguistic Theory*, 5(2):201–239.
- Kunter, G. (2011). *Compound stress in English. The phonetics and phonology of prosodic prominence*. De Gruyter, Berlin.
- Kunter, G. and Plag, I. (2007). What is compound stress? In *Proceedings of the 16th International Congress of Phonetic Sciences*, pp. 1005–1008, Saarbrücken.
- Kvam, A. M. (1990). Three-part noun combinations in English. Composition - meaning - stress. *English Studies*, 71:152–160.
- Ladd, R. (1984). English compound stress. In Gibbon, D. and Richter, H. (eds.), *Intonation, accent and rhythm*, pp. 253–266. De Gruyter, Berlin.
- Ladd, R. D. (1996). *Intonational Phonology*. Cambridge University Press, Cambridge.
- Ladefoged, P. (2003). *Phonetic data analysis*. Blackwell, Malden, MA.
- Ladefoged, P. (2006). *A course in Phonetics*. Thomson Wadsworth, Boston, MA, 5th edition.
- Lappe, S. and Plag, I. (2007). The variability of compound stress in English: towards an exemplar-based alternative of the compound stress rule. In *Proceedings of the ESSLLI workshop on exemplar-based models of language acquisition and use*, pp. 8–23, Dublin. Trinity College.
- Lehiste, I. (1970). *Suprasegmentals*. MIT Press, Cambridge, MA.
- Liberman, M. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2):249–336.
- Liberman, M. and Sproat, R. (1992). The stress and structure of modified noun-phrases in English. In Sag, I. A. and Szabolcsi, A. (eds.), *Lexical Matters*, pp. 131–181. Center for the Study of Language and Information, Stanford.
- Marchand, H. (1969). *The categories and types of present-day English word formation: a synchronic-diachronic approach*. C. H. Beck'sche Verlagsbuchhandlung, München, 2nd edition.

- Meyer, R. (1993). *Compound Comprehension in Isolation and in Context. The Contribution of Conceptual and Discourse Knowledge to the Comprehension of Novel Noun-Noun Compounds*. Niemeyer, Tübingen.
- Neppert, J. and Pétursson, M. (1986). *Elemente einer akustischen Phonetik*. Buske, Hamburg, 2nd edition.
- Olsen, S. (2000). Compounding and stress in English: A closer look at the boundary between morphology and syntax. *Linguistische Berichte*, 181:55–69.
- Olsen, S. (2001). Copulative compounds: A closer look at the interface between syntax and morphology. In Booij, G. . J. v. M. e. (ed.), *Yearbook of morphology 2000*, pp. 279–320. Kluwer, Dodrecht / Boston/London.
- Ostendorf, M., Price, P., and Shattuck-Hufnagel, S. (1996). *Boston University Radio Speech Corpus*. Linguistic Data Consortium, Philadelphia, PA.
- Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *Journal of the Acoustic Society of America*, 66:362–370.
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. PhD thesis, Massachusetts Institute of Technology.
- Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-Effects-Models in S and S-Plus*. Springer, New York.
- Plag, I. (2003). *Word-Formation in English*. Cambridge University Press, Cambridge.
- Plag, I. (2006). The variability of compound stress in English: structural, semantic, and analogical factors. *English Language and Linguistics*, 10(1):143–172.
- Plag, I. (2010). Compound stress assignment by analogy: the constituent family bias. *Zeitschrift für Sprachwissenschaft*, 29(2):243–282.
- Plag, I. and Kunter, G. (2010). Constituent family size and compound stress assignment in English. In Olsen, S. (ed.), *New Impulses in Word-Formation*, number 17 in *Linguistische Berichte*, pp. 349–382. Buske, Hamburg.
- Plag, I., Kunter, G., and Lappe, S. (2007). Testing hypotheses about compound stress assignment in English: a corpus-based investigation. *Corpus Linguistics and Linguistic Theory*, 3(2):199–233.

- Plag, I., Kunter, G., Lappe, S., and Braun, M. (2008). The role of semantics, argument structure, and lexicalization in compound stress assignment in English. *Language*, 84(4):760–794.
- Plag, I., Kunter, G., and Schramm, M. (2011). Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics*, 39:362–374.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Schmerling, S. F. (1971). A stress mess. *Studies in the Linguistics Science*, 1:55–66.
- Selkirk, E. (1984). *Phonology and Syntax: the relation between sound and structure*. MIT Press, Cambridge, MA.
- Sepp, M. (2006). *Phonological constraints and free variation in compounding: a corpus study of English and Estonian noun compounds*. PhD thesis, City University, New York.
- Shattuck-Hufnagel, S., Ostendorf, M., and Ross, K. (1994). Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics*, 22:357–388.
- Spencer, A. (1996). *Phonology*. Blackwell, Oxford.
- Spencer, A. (2003). Does English have productive compounding? In Booij, G. E., DeCesaris, J., Ralli, A., and Scalise, S. (eds.), *Topics in Morphology. Selected papers from the 3rd mediterranean morphology meeting*, pp. 329–341. Institut Universitari de Lingüística Aplicada, Barcelona.
- Sproat, R. (1994). English noun-phrase accent prediction for text-to-speech. *Computer Speech and Language*, 8:79–94.
- Summers, D. (2000). *Longman Advanced American Dictionary*. Longman, Harlow, Essex.
- Teschner, R. and Whiteley, M. (2004). *Pronouncing English*. Georgetown University Press.
- Visch, E. (1999). The rhythmic organization of compounds and phrases. In Van der Hulst, H. (ed.), *Word Prosodic systems in the Languages of Europe*, pp. 161–221. Mouton de Gruyter.

Warren, B. (1978). *Semantic patterns of noun-noun compounds*. Acta Universitatis Gothoburgensis, Göteborg.

Whalen, D. H. and Levitt, A. (1995). The universality of intrinsic f_0 of vowels. *Journal of Phonetics*, 23:349–366.