

Linguistische Arbeiten

508

Herausgegeben von Peter Blumenthal, Gereon Müller,
Ingo Plag, Beatrice Primus, Klaus von Heusinger und Richard Wiese

Stefan Baumann

The Intonation of Givenness

Evidence from German

Max Niemeyer Verlag
Tübingen 2006



Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN 13 978-3-484-30508-3 ISSN 0344-6727

ISBN 10 3-484-30508-8

© Max Niemeyer Verlag, Tübingen 2006

Ein Unternehmen der K.G. Saur Verlag GmbH, München

<http://www.niemeyer.de>

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Printed in Germany.

Gedruckt auf alterungsbeständigem Papier.

Druck: Laupp & Göbel GmbH, Nehren

Einband: Nadele Verlags- und Industriebuchbinderei, Nehren

The Intonation of Givenness

Evidence from German

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Philosophie

der Philosophischen Fakultäten der Universität des Saarlandes

vorgelegt von

Stefan Baumann

aus Kiel

Dekan: Prof. Dr. Wolfgang Haubrichs
Berichterstatter/innen: Prof. Dr. Martine Grice
Prof. Dr. William J. Barry
Prof. D. Robert Ladd (Univ. of Edinburgh)
Tag der letzten Prüfungsleistung: 15. 12. 2005

Für meine Eltern
Elisabeth und Ernst-Otto Baumann

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der intonatorischen Markierung von Informationsstruktur, insbesondere auf der Ebene der kognitiven Aktivierungsgrade von Diskursreferenten. Ein übergeordnetes Ziel besteht darin, Intonation und die ‘Gegebenheit’ (*Givenness*) referierender Ausdrücke miteinander zu verbinden. Diese beiden Gebiete gelten oftmals aus Sicht des jeweils anderen Gebiets als entweder zu vage oder zu komplex, um sie in die eigene Forschung zu integrieren. Die Ergebnisse der vorliegenden empirischen Untersuchung beziehen sich auf das Deutsche, sind aber größtenteils auf andere westgermanische Sprachen wie das Englische und Niederländische übertragbar.

Nach einer Einführung in die Ziele und Struktur der Arbeit (Kapitel 1) wird in Kapitel 2 der theoretische Rahmen der Untersuchung abgesteckt. Hierbei werden zunächst die phonetischen Aspekte der Intonation (mit dem Schwerpunkt Akzentuierung) sowie die hier zugrunde gelegte phonologische Theorie, die Autosegmental-Metrische Phonologie, vorgestellt. In diesem Zusammenhang erfolgt außerdem die Vorstellung von GToBI, dem in dieser Arbeit verwendeten Modell zur Annotation deutscher Intonationsmuster.

Anschließend wird, weitgehend theorie-neutral, in das Konzept der Gegebenheit eingeführt und eine Abgrenzung zu anderen Dimensionen der Informationsstruktur, wie der Fokus-Hintergrund- und Thema-Rhema-Gliederung, vorgenommen. Ferner wird diskutiert, auf welche Bereiche und Konstituenten sich Gegebenheit (und Neuheit) bezieht, wodurch sie hervorgerufen wird und welche Rolle dem Sprecher und Hörer bei ihrer Bestimmung zukommt. Es werden drei (kognitive) Ebenen unterschieden, die in der Literatur mit dem Konzept der Gegebenheit in Beziehung gesetzt werden, nämlich die Ebenen ‘Wissen’, ‘Bewußtsein’ und ‘Wichtigkeit für den Sprecher’. In Anlehnung an Lambrecht (1994) betrachte ich die ersten beiden Ebenen als zentral für meine Definition von Gegebenheit, weil sie sich auf den kognitiven Status von Diskursreferenten oder Propositionen konzentrieren. Sie entsprechen den Ebenen der ‘Identifizierbarkeit’ und ‘Aktivierung’. Erstere beschreibt die Fähigkeit des Hörers, aus der Menge aller möglichen Referenten, die durch einen bestimmten sprachlichen Ausdruck kodiert werden, denjenigen auszuwählen, den der Sprecher gemeint hat. Letztere kennzeichnet den Grad, zu dem ein Referent oder eine Proposition im Moment der Äußerung im Bewußtsein des Hörers abrufbar ist. Die dritte Ebene bezeichnet die pragmatische Rolle eines Diskursreferenten in einer Proposition, ausgedrückt durch die Unterscheidung zwischen Fokus (was dem Sprecher wichtig ist) und Hintergrund (was dem Sprecher nicht wichtig ist).

Auf Grundlage der Differenzierung dieser Ebenen werden in einem nächsten Schritt Fragen ihrer sprachliche Realisierung diskutiert. Dabei gilt der linguistischen

Markierung von Gegebenheit im engeren Sinne, d.h. der Ebenen von Identifizierbarkeit und Aktivierung, ein besonderes Interesse. Es zeigt sich, dass (Nicht-)Identifizierbarkeit vor allem durch (In-)Definitheit gekennzeichnet ist, während Aktivierung, ein Parameter, der die Identifizierbarkeit eines Referenten voraussetzt, durch zwei verschiedene linguistische Beschreibungsebenen kodiert wird: lexikalische Form und Intonation. Bereits aktivierte Referenten werden oft als Pronomen realisiert, während weniger aktive Konzepte im Allgemeinen in ihrer vollen lexikalischen Form auftreten. Ferner sind aktivierte Referenten oftmals unakzentuiert, wohingegen halb- oder nicht-aktivierte Referenten in aller Regel einen Akzent tragen.

Allerdings zeigt diese Studie, dass die Korrelation der Dichotomien ‘akzentuiert versus unakzentuiert’ und ‘neu versus gegeben’ eine unzulässige Vereinfachung darstellt. Viele neuere Arbeiten über den kognitiven Status von Diskursreferenten betrachten die Unterscheidung von neuer und gegebener Information nicht mehr als binär, sondern als graduell. Ein solches Kontinuum möglicher Aktivierungsgrade läßt sich allerdings sprachlich nicht angemessen widerspiegeln, weil die Anzahl der zur Verfügung stehenden linguistischen Kategorien begrenzt ist. Ich beschränke mich daher auf die Untersuchung dreier verschiedener kognitiver Zustände, die ich in Anlehnung an Chafe (1994) ‘gegeben’ (*Given*), ‘erschließbar’ (*Accessible*) und ‘neu’ (*New*) nenne und von denen ich annehme, dass sie durch distinktive formale, insbesondere prosodische Kategorien markiert werden.

Der bislang kaum untersuchte Bereich der ‘erschließbaren’ bzw. ‘halb-aktivierten’ Information und seiner sprachlichen Realisierung ist Hauptgegenstand meiner Studie. Während die morphosyntaktische Markierung erschließbarer Referenten relativ unumstritten ist (sie treten zumeist in Form definiter Nominalphrasen auf), gibt es keinen Konsens bezüglich ihrer prosodischen Markierung. Dies ist unter anderem darauf zurückzuführen, dass in den Beispielen, die in der einschlägigen Literatur diskutiert werden, die Position der halb-aktivierten Referenten innerhalb der Äußerung variiert. Folglich wird der Unterschied zwischen prä nuklear und nuklear akzentuierten Ausdrücken vernachlässigt. Der Status prä nuklearer Akzente ist allerdings weit weniger eindeutig als derjenige nuklearer Akzente, nicht zuletzt weil die Verteilung und Stärke prä nuklearer Akzente in größerem Maße von der rhythmischen Struktur und Länge der Äußerung abhängt. Ich habe mich daher auf die prosodische Markierung des letzten – und somit potentiell nuklearen – Arguments in deklarativen und vollständig fokussierten Äußerungen konzentriert.

Auf diesem theoretischen Fundament stehen die empirischen Untersuchungen, die die Analyse eines gelesenen Korpus deutscher Nachrichtentexte (Kapitel 3) sowie zwei Perzeptionsexperimente (Kapitel 4) umfassen. Die Korpusanalyse, die zur Erhebung von Produktionsdaten dient, ist dabei lediglich als Vorstufe der Untersuchung zu betrachten, weil die Daten von einer einzigen Sprecherin gelesen wurden und somit die Repräsentativität der Intonationsmuster nicht uneingeschränkt

gewährleistet werden kann. Dennoch können gewisse Korrelationen zwischen Akzenttyp und kognitivem Aktivierungsgrad von Referenten gefunden werden, die anschließend in kontrollierten Perzeptionsexperimenten überprüft werden. In diesen Experimenten haben die Versuchspersonen die Aufgabe, die Angemessenheit bestimmter Akzenttypen (H^* , $H+L^*$) sowie Deakzentuierung als Kennzeichen eines Referenten in verschiedenen Kontexten zu beurteilen. Die Auswahl der getesteten Akzenttypen basiert – neben den Daten der Korpusanalyse – auf Vorschlägen aus früheren Studien zum Englischen und Deutschen, vor allem von Pierrehumbert & Hirschberg (1990) und Kohler (1991a). Das erste Experiment hat die Untersuchung von Akzenttyp-Präferenzen auf Referenten in unterschiedlichen Priming-Bedingungen zum Gegenstand, d.h. nach vorangehender auditiver, visueller oder fehlender Darbietung des gleichen Referenten. Dem liegt die Annahme zugrunde, dass der unterschiedliche Präsentationsmodus die Referenten in unterschiedlichem Maße aktiviert, so dass den drei Bedingungen gegebene, erschließbare und neue Information entspricht. Das zweite Experiment konzentriert sich weitgehend auf die Prosodie erschließbarer Referenten, insbesondere verschiedener Arten textuell und inferentiell erschließbarer Information – im Gegensatz zum ersten Experiment, in dem situativ (d.h. visuell) erschließbare Referenten dargeboten wurden. Es werden die gleichen Akzenttypen (inklusive Deakzentuierung) verwendet wie im ersten Experiment.

Die Ergebnisse der empirischen Untersuchungen bilden die Grundlage des hier vorgeschlagenen Modells von Intonation und Gegebenheit im Deutschen (Kapitel 5). Im einzelnen zeigt sich, dass Neuheit mit dem Akzenttyp H^* korreliert. Dies gilt sowohl für ‘brandneue’, d.h. für den Hörer nicht identifizierbare Referenten als auch für ‘ungebrauchte’, d.h. dem Hörer zwar bekannte, aber aus dem vorangehenden Diskurs nicht ableitbare (und damit nicht aktivierte) Referenten. Der formale Unterschied zwischen diesen beiden Kategorien besteht in ihrer morphosyntaktischen Markierung: Brandneue Information wird als indefinit, ungebrauchte als definit gekennzeichnet. Am anderen Ende der Skala läßt sich feststellen, dass gegebene, d.h. vollständig aktivierte Information gemeinhin nicht akzentuiert wird. Die intonatorische Markierung von Referenten im Zwischenbereich der ‘erschließbaren’ Information ist hingegen weniger eindeutig, auch wenn eine Präferenz von Akzenttyp $H+L^*$ für bestimmte Arten halb-aktivierter Information, wie z.B. ein durch ein ‘Szenario’ mit-aufgerufenes Konzept (z.B. *Restaurant – Kellner*) oder die Anapher in einer Ganzes-Teil-Relation (z.B. *Buch – Seiten*), nachgewiesen werden kann. Andere halb-aktivierte Referenten, wie Synonyme (z.B. *Apfelsine – Orange*) oder die Anapher in einer Teil-Ganzes-Relation (z.B. *Seiten – Buch*) oder einer Äquivalenz-Relation (z.B. *A=Bs Schwester $\approx$ B=As Bruder*), werden bevorzugt deakzentuiert.

Generell zeigt sich, dass Information zwischen den Polen ‘gegeben’ und ‘neu’ im Hinblick auf ihre intonatorische Markierung nicht klar abgrenzbar ist und dass

unterschiedliche Arten mehr oder weniger aktivierter Information, die z.B. verschiedene lexikalisch-semantische Relationen kodieren, unterschiedliche Akzenttypen zu ihrer Kennzeichnung erfordern. Dabei deutet sich eine (pseudo-graduelle) Rangfolge der Akzenttypen an – wobei die Tonhöhe auf der lexikalisch betonten Silbe des referierenden Ausdrucks den entscheidenden Faktor darstellt –, die mit Änderungen im Aktivierungsgrad des Referenten einhergeht. Dieser Zusammenhang suggeriert eine gewissermaßen ‘ikonische’ Funktion der Tonhöhe, die sich auch in Gussenhovens (2002) *Effort Code* widerspiegelt: je höher der Ton auf der lexikalisch starken Silbe, desto ‘neuer’ (oder wichtiger) der Diskursreferent. Diese Faustregel erklärt auch den Zwischenstatus des Akzenttyps H+L*, der in allen Fällen, in denen einer der beiden anderen getesteten Markierungen (H* oder kein Akzent) bevorzugt wird, als zweitbeste Möglichkeit eingestuft wird.

Diese Rangordnung verschiedener Akzenttypen beinhaltet auch eine Abstufung in der wahrgenommenen Akzentstärke. Mehrere Studien zum Deutschen und Englischen postulieren mehr oder weniger direkt die Relevanz eines ‘Sekundärakzents’ als Kennzeichen halb-aktivierter Information. Allerdings können Art, Position und genaue Funktion dieser sekundären Prominenzen variieren, wie etwa zwischen ‘Sekundärakzenten’ im Sinne Bürings (Büring 2003), ‘Dauerakzenten’ (Kohler 2005) oder ‘Phrasenakzenten’ (Grice, Ladd & Arvaniti 2000).

Die Ergebnisse dieser Arbeit sind auch für sprachtechnologische Anwendungen relevant, vor allem im Bereich der Sprachsynthese. Die Integration eines feiner differenzierten Intonationsmoduls kann die Natürlichkeit eines Synthesesystems erheblich steigern, insbesondere wenn der Zugriff auf den sprachlichen Kontext sowie auf semantische Informationen, die für die korrekte Zuweisung der verschiedenen Akzenttypen (inklusive Deakzentuierung) von entscheidender Bedeutung sind, gewährleistet ist. Dies kann durch die Verknüpfung des Synthesesystems mit einer semantischen Datenbank wie GermaNet erreicht werden. Ein solches System wäre nicht mehr ausschließlich textbasiert – und seine Akzentzuweisungsregeln nicht mehr allein von Wortart-Informationen abhängig – sondern die Vorstufe einer *concept-to-speech*-Synthese mit kontext-sensitiver Intonation.

Acknowledgements

Many people contributed to the completion of this thesis.

Firstly, I would like to thank my supervisor Martine Grice for her invaluable support and assistance throughout all stages of this study, for her enthusiasm and countless inspiring discussions.

I am also deeply grateful to my second supervisor and mentor, Bill Barry, who was always there when I needed advice, and to Bob Ladd, who kindly agreed to write an external review. This meant a lot to me since it was one of Bob's courses in Edinburgh several years ago that aroused my interest in the subject.

Furthermore, I feel indebted to Richard Wiese, editor of the series ‚Linguistische Arbeiten‘, for his helpful comments on the manuscript.

Also thanks to the many friends and colleagues who helped me in various ways, with both intellectual as well as practical contributions, especially in setting up and running the experiments. In Saarbrücken, these people are (among others) Kerstin Hadelich, Caren Brinckmann, Gudrun Schuchmann, Bettina Braun, Jürgen Trouvain, Marc Schröder, Bistra Andreeva, Cordula Klein, Silke Jarmut, Andrea Weber, Jacques Koreman, Thomas Blug, Anja Moos, Kerstin Kunz, Stephanie Becker and Silvia Hansen-Schirra. In Cologne, I would like to cordially thank the whole team, namely Doris Mücke, Maja Warnking, Philipp von Böselager, Barbara Hugo-Dilworth, Johannes Becker, Anne Hermes, Jacqueline Anthes, Kyung-Hee Kim, Anna Diagne, Christine Riek and Theo Klinker.

And finally, special thanks to my Schatje, Stella Neumann, who supported me in ALL matters.

Contents

1	Introduction.....	1
1.1	Motivation and Aims	1
1.2	Structure of the Study.....	2
2	Theoretical Background	5
2.1	Intonation	5
2.1.1	Phonetics of Intonation	5
2.1.2	Phonology of Intonation	12
2.1.2.1	Principles of Autosegmental and Metrical Phonology.....	12
2.1.2.2	Intonation in Autosegmental-Metrical Phonology	18
2.1.2.3	GToBI	23
2.2	Givenness	28
2.2.1	Givenness and Information Structure	28
2.2.2	Domains and Modes of Givenness.....	39
2.2.3	Perspectives of Givenness	43
2.2.4	Levels of Givenness.....	48
2.2.5	Degrees of Givenness and their Linguistic Marking	57
2.2.5.1	The Marking of (Non-)Identifiability	58
2.2.5.2	The Marking of (In-)Activation.....	65
2.2.5.3	The Marking of Focus-Background Structure	89
2.3	Intermediate Summary: The Relation between Intonation and Givenness	96
3	Corpus Analysis	100
3.1	The MULI Database.....	100
3.2	Analysis	103
3.3	Discussion.....	109
4	Experiments	114

4.1 Perception Experiment I: Accent Type and Modes of Givenness	114
4.1.1 Motivation.....	114
4.1.2 Hypotheses.....	116
4.1.3 Design of Test Material.....	116
4.1.3.1 Modes and Degrees of Givenness Investigated.....	116
4.1.3.2 Visual Test Material.....	117
4.1.3.3 Auditory Test Material	118
4.1.4 Experimental Setup	121
4.1.5 Results.....	123
4.1.6 Discussion.....	124
4.2 Perception Experiment II: Accent Type and Types of Accessibility	126
4.2.1 Motivation.....	126
4.2.2 Hypotheses.....	126
4.2.3 Design of Test Material.....	126
4.2.3.1 Types of Accessibility Investigated	126
4.2.3.2 Textual Test Material.....	130
4.2.3.3 Auditory Test Material	133
4.2.4 Experimental Setup	135
4.2.5 Results.....	135
4.2.6 Discussion.....	136
4.2.7 Summary and Conclusion.....	141
4.2.8 Digression: Intonation, Biological Codes and their Linguistic Manifestations	143
5 A Model of Intonation and Givenness	150
6 Summary and Outlook	162
Bibliography	167

1 Introduction

1.1 Motivation and Aims

This study is intended to address two kinds of audience: intonologists who are interested in information structure, and experts in the semantics and pragmatics of discourse who have an interest in prosodic aspects of information structure. By addressing both, we aim to forge links between the two fields, taking into account the growing demand articulated in recent years.

In order to establish a ‘common ground’, an overview of the state of the art in both intonation research and information structure will be given, the latter predominantly with respect to the cognitive states of discourse referents, subsumed under the notion of ‘Givenness’. Since the notion of Givenness has been used in the literature in diverging ways, applying to different levels of description, we will distinguish it from other levels of information structure and propose a model of ‘Givenness proper’ and its linguistic marking.

Multiple functions can be attributed to intonation, ranging from the clearly paralinguistic encoding of emotions, over ‘more linguistic’ pragmatic functions such as indicating speech act distinctions, to expressing strictly linguistic contrasts at word level in tone languages. We will concentrate on the (linguistic) functions of intonation which serve to assign a structure to utterances in terms of phrasing and prominence relations and which are relevant for information packaging. This structure, which is often realised with a combination of phonetic properties (predominantly perceived pitch, but also entailing loudness, vowel quality, and relative length of syllables and words) is determined phonologically in terms of abstract tonal values at the edges of phrases and on prominent syllables. Of particular interest is the distribution of prominences or (pitch) accents, which fulfil the highlighting function of intonation.

In studies on the realisation of information structure in West Germanic languages (notably English, German and Dutch), it is commonly assumed that important or New information is marked by a pitch accent, while unimportant or Given information is either unaccented or deaccented (i.e. there is no pitch accent where one would otherwise be expected). A central aim of this study is to show that this view is an inadequate simplification. This applies to both fields of research:

First, there is evidence from psycholinguistically oriented studies that the difference in the cognitive activation of discourse referents is gradual in nature and – consequently – that there are ‘degrees of Givenness’ between the extreme poles of Given and New. Most linguistic studies which acknowledge the existence of Givenness degrees are predominantly concerned with the morphosyntactic form of

referring expressions, since they usually investigate written language. Only few approaches combine morphosyntax and intonation.

Second, the intonational means used for the encoding of information go far beyond a dichotomy of accent versus lack of accent. As a number of investigations of the prosodic marking of information structure have shown, different types of accent have to be taken into account, as well as different degrees in the strength of accents.

Thus, it will be investigated in this study how far not only accentuation and lack thereof, but also type of accentuation can be used to indicate different degrees of Givenness, and in particular how they are used in German. We will gain our evidence from empirical data, elicited in both production and perception experiments. Almost all – of the few – empirical studies on the relation between activation degrees and intonation that have been conducted so far are based on English data. Since German and English are said to be closely related, at least in terms of their intonational systems, being both West Germanic (see Ladd 1996), the claims made in the literature on English are taken as a point of departure for our own investigation. Having said this, however, we should be aware that structures or relations that hold for one language cannot be automatically transferred to another language, however related. We therefore stress that the model of intonation and Givenness proposed here applies to German and although it may be relevant for research into English, it is not claimed that the results hold for English.

1.2 Structure of the Study

Since we are bringing together two only partially intersecting fields of research, we shall first give separate accounts of their theoretical background (chapter 2). Section 2.1 deals with intonation. First, phonetic aspects of intonation (understood as being equivalent to the more general term ‘prosody’) will be discussed, in particular the phonetic parameters which are relevant for the description of speech melody, phrasing and – most importantly – accentuation (section 2.1.1).

We shall then turn to the phonological description of intonation (section 2.1.2), starting with a brief overview of the principles of Autosegmental and Metrical Phonology (section 2.1.2.1). The combined autosegmental-metrical theory is the currently most widespread framework for representing pitch accents and prosodic phrasing and the one used in this study (presented in section 2.1.2.2). Finally, the annotation system GToBI (German Tones and Break Indices) will be introduced, which has been developed for the description of German intonation within the autosegmental-metrical framework (section 2.1.2.3). This system will be used for the annotation of examples throughout the whole study.

Section 2.2 deals with aspects centering around the notion of Givenness. We will first develop a definition of the concept of Givenness and other basic information

structural concepts. Then the role of Givenness within the vast field of information structure will be determined, delimiting it in particular from the dimensions of background versus focus and theme versus rheme (section 2.2.1). This means that Givenness proper, in our understanding, exclusively applies to the status of referents or propositions in discourse, not to the partitioning of a sentence or utterance.

The following subsection (2.2.2) deals with the domains and modes of constituents that can be called Given or New, i.e. with questions of their nature (e.g. referents or propositions), size (e.g. one item or a whole phrase) and origin (e.g. recoverable from the preceding discourse or the physical context). Furthermore, the role of speaker and listener in determining what is Given and what is New (section 2.2.3) will be discussed. Finally, in sections 2.2.4 and 2.2.5, three levels are differentiated to which the notion of Givenness has been applied in the literature. These are the levels of knowledge, consciousness and newsworthiness, which will be explained in detail. Especially the second level of consciousness or activation, which represents the core of our understanding of Givenness, allows for a differentiation of various degrees of Givenness. We will propose to add at least one intermediate state between the extreme poles of Given and New, namely Accessibility, and give a detailed account of how these different degrees or states of cognitive activation are marked linguistically, i.e. by morphosyntactic and – in particular – prosodic means.

As a first step towards a comprehensive model of intonation and Givenness in German, a read corpus of German newspaper texts, the MULI corpus, was analysed (chapter 3). Since the corpus was read by a single speaker, the attested relationship between activation states of discourse referents and their prosodic marking can only be regarded as tendencies which should be verified in further experiments involving a number of speakers or in perception experiments involving a (large) number of listeners.

We carried out the latter, described in chapter 4. In two perception experiments, listeners judge the appropriateness of the presence or absence of accentuation, as well as accent type, as a marker of discourse referents in various contexts. The first experiment investigates preferences for accent type and placement across different conditions: where, prior to the target utterance, the referent has been presented in the form of a picture (visual priming) or in spoken form (auditory priming), and where it has not been presented at all (no priming). The different modes of presentation are assumed to activate the referents to different degrees, thus eliciting a distinction between Given, Accessible and New information (section 4.1).

The second experiment concentrates on the intonation of the largely unstudied area of information between the poles Given and New. Using the same accent types as in the first experiment, various kinds of textual and inferential Accessibility are examined, again assuming differences in activation, this time brought about by differences in the semantic relation between an antecedent and an anaphor (section 4.2). The discussion of the experimental results leads to a brief discussion of the

iconicity of intonation and the question to what extent intonational meaning is determined by biological codes (section 4.2.8).

Based on our empirical data, and taking into account several aspects of previous approaches, we propose in chapter 5 a fine-grained model of intonation and Givenness for German, which overcomes a simple binary distinction between New and Given on the one hand and accented versus unaccented on the other.

Finally, in chapter 6, the main results are summarised, and suggestions are put forward for future research and as to how parts of the proposed model can be made use of in technological applications, in particular in speech synthesis.

2 Theoretical Background

2.1 Intonation

In spoken language, intonation serves diverse linguistic and paralinguistic functions, ranging from the marking of sentence modality to the expression of emotional and attitudinal nuances. However, the term ‘intonation’ has been defined in at least two different ways in the literature. A narrow definition equates intonation with ‘speech melody’, restricting it to the “ensemble of pitch variations in the course of an utterance” (t Hart et al. 1990: 10). A broader account of intonation, which will be adopted for the present study, is equivalent to what is often called ‘prosody’, subsuming such different phenomena as pitch movements and range (speech melody), the division of speech into chunks (phrasing), highlighting at word level (lexical stress) and utterance level (accentuation), the marking of prominence relations (rhythm) and variations in speech rate (tempo).

All these phenomena, most of which serve to assign a phonological structure to utterances, are realised with a combination of phonetic properties such as perceived pitch, loudness, vowel quality, and relative length of syllables, words and pauses. These properties, especially the phonetic correlates of accentuation, which are of particular interest for this study, will be dealt with in section 2.1.1. Basics of the phonological theory adopted here will be presented in section 2.1.2, including a brief description of the principles of Autosegmental and Metrical Phonology (section 2.1.2.1), the representation of intonation within this framework (section 2.1.2.2) and, finally, a description of the annotation model for German intonation used in the present study (section 2.1.2.3).

2.1.1 Phonetics of Intonation

The phonetic parameter that is most centrally associated with intonation – especially in the narrow sense of ‘speech melody’ – is *pitch*. However, the term ‘pitch’, or, more precisely, ‘variations of pitch’, only denotes the perceptual impression of speech melody, which has physiological and acoustic correlates. The physiological or articulatory source of pitch variations are changes in the rate at which the vocal folds vibrate. The vibration is a result of aerodynamic forces acting upon the folds after they have been sufficiently approximated by some of the laryngeal muscles. Acoustically, the frequency at which the vocal folds vibrate correlates with the fundamental frequency or F0. It denotes the repetition frequency of a complex, quasi-periodic sound wave, which is equivalent to the highest common factor of the sound wave’s component frequencies or ‘harmonics’ (see Ladefoged 1962: 111). The fundamental frequency is measured in ‘Hertz’ (Hz), replacing the older term ‘cycles

per second' (cps), which directly mirrors the cyclic opening and closing of the glottis. The higher the frequency of vocal fold vibrations and, in turn, the higher the fundamental frequency of a sound, the higher is its perceived pitch. Normally, we are able to hear sounds between 40 and 4000 Hz (see 't Hart et al. 1990: 26), with normal speaking voices ranging between 150 and 400 Hz for women and between 80 and 200 Hz for men. However, pitch values can be perceived even if the F0 is missing, as e.g. in hissed (i.e. voiceless) sounds or in conversations on the telephone, which does not transmit sound waves below 300 Hz. In these cases, the human ear can detect the fundamental frequency by measuring the difference between the harmonics (which are integer multiples of F0).

The rate at which the vocal folds vibrate is determined by their elasticity, length and mass (with vocal folds in females being shorter and lighter than in males, thus producing higher vibration frequencies) but also, more indirectly, by muscular tension and the amount of air pressure below the glottis. Increased tension stretches the vocal folds which, in turn, thins their effective vibrating portion and leads to higher F0 values. Enhanced subglottal pressure, on the other hand, primarily increases the *amplitude* of vocal folds vibration, not its frequency. Nevertheless, a higher amplitude induces a greater deformation of the folds, which, in turn, leads to greater mechanical stiffness and – as a consequence – faster vibration of the vocal folds (see 't Hart et al. 1990: 13f.).

While the two influencing factors on F0 just mentioned, i.e. muscular tension and sub-glottal air pressure, are to a large extent under the speaker's control (see Borden & Harris 1984: 74ff.), other physiological factors, like certain supralaryngeal articulatory gestures, are not. For example, high vowels have higher intrinsic pitch than low vowels (especially in prominent syllables; see e.g. Lehiste & Peterson 1961, Ladd & Silverman 1984, Di Cristo & Hirst 1986), and the F0 of vowels in general is affected by adjacent consonants. A voiceless obstruent, for instance, leads to a higher F0 at the beginning of the following vowel than a voiced obstruent due to increased vocal fold tension after a glottalic opening gesture (see Kingston 1991, Gussenhoven 2004). Such minor perturbations in the F0 curve do not influence listeners' interpretation of the intonation contour (see Silverman 1987).¹ They are known as instances of 'microprosody' or 'microintonation'.

More complicated than determining the phonetic correlate of speech melody is the answer to the question of what constitutes *accent*. Among others, Kohler (1977) and Beckman (1986) claim for German and (American) English, respectively, that the acoustic correlate of accentuation is a complex mixture of F0 variation, increased intensity (primarily equivalent to higher subglottal pressure) and increased duration

¹ Compare, however, Kingston & Diehl (1994), who argue that the phenomena just mentioned may be controlled articulations which are intended to enhance height or voicing contrasts. Nevertheless, they are not controlled for intonation purposes.