Judith Haan

Speaking of Questions

An Exploration of Dutch Question Intonation

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6000
e-mail: lot@let.uu.nl
http://www.let.uu.nl/LOT/

Cover illustration:
Medieval rule for the recitation of liturgical texts (Münster), taken from H. Helmholtz, *Die Lehre von Tonempfindungen*, Braunschweig 1877. For each of the punctuation marks the rule prescribed the appropriate inflection of the voice, corresponding with the intonation of actual speech. Thus, a comma required a minor final rise (‘sic canta comma,’), a colon a minor final fall (‘sic duo puncta:’), a full stop a substantial final fall (‘sic vero punctum.’), and the interrogation mark final rising pitch (‘Sic signum interrogationis?’).

ISBN 90-76864-13-6
NUGI 941

This research was funded in part by the Netherlands Organsation for Research (NOW) under grant 200-50-073.

Table of Contents

3.
PRODUCTION (1):
QUESTION INTONATION IN READ SPEECH

4.
PRODUCTION (2):
QUESTIONS AND ACCENTUATION

7.
SUMMARY AND CONCLUSIONS

Chapter 1

Introduction and background (1): Interrogativity

1.1 Introduction

1.1.1 Preliminaries

If speech serves the goal of human communication, the speech act of questioning does so *in optima forma*. More than any other act performed by speech, a question draws the addressee into interaction with the speaker. The addressee is expected to respond, commonly by agreeing or disagreeing, or by filling a specific blank in the speaker's knowledge. An addressee who fails to meet the speaker's expectation for some response is felt to break an elementary communicative rule. Thus, asking questions is a highly suitable means for starting and keeping up all kinds of interactions, such as conversations and interviews.

Primarily, however, questioning is an effective means of making people communicate information that is relevant to the inquirer. Young children learn about life by continually asking questions. In adults, questioning may be inspired by various motives, ranging from a speaker's eagerness to expand his knowledge, to the more down-to-earth desire to satisfy one's curiosity, or to obtain some trivial piece of information such as the whereabouts of the post office. The present study focusses on questioning in so far as it serves these cognitive needs of the speaker[1]. One might wonder whether this goal cannot be attained in a more parsimonious way, without specific interrogative strategies. In theory, the declarative sentence *I would like you to specify the period in which the pyramids were built*, or the imperative sentence *Specify the period in which the pyramids were built* would appear to have the same function as the question *When were the pyramids built?*. Nonetheless, languages lacking explicit lexical, syntactic and/or prosodic strategies for encoding interrogativity seem to be non-existent (cf. Chisholm 1982:278). It seems justified, therefore, to regard question utterances as a grammatical category in its own right and, consequently, as a proper object for linguistic study.

[1] If, today, questioning is generally taken to serve the purpose of expanding a speaker's actual state of knowledge, past times have also seen a claim to the opposite. The philosopher Socrates insisted that people could obtain true knowledge, not by being taught or by asking questions, but by *being* interrogated. That is, a skilled teacher was assumed to be able to guide a pupil towards knowledge that was already dormant in the latter's soul, by systematically asking him the right questions (cf. Plato, in *Meno)*.

1.1.2 Research questions

The present study concerns itself with one particular aspect of questions, i.e. their intonation, in one particular language, i.e. Dutch[2]. The central research questions can be formulated as follows:

- Does the discoursal function 'interrogativity' systematically correspond with specific intonational properties?
- If this is the case, are these properties phonological, phonetic, or both?

In the event of systematic correspondences, two further research questions present themselves. First, do the observed intonational properties of interrogativity vary as a function of the *type* of question they occur in? In fact, it would seem plausible for such intonational properties to be stronger as other markers of interrogativity (e.g. the presence of a question word, or the inversion of subject and finite verb) are absent. Second, do the observed intonational properties vary as a function of the speaker's sex? Empirical studies have shown (i) that female speech, apart from being higher-pitched for anatomical or cultural reasons, is more expressive, more varied and more listener-directed than male speech, (ii) that female communicative behaviour is primarily directed towards interaction whereas that of men aims at assertion and carrying out tasks, and (iii) that women are more prepared to show the dependence implied by questioning than are men. Considering that interrogativity aims at establishing interaction with a listener and that questioning gives expression to a speaker's dependence, we expect intonational properties of interrogativity to be more pronounced in female speech.

Before it is possible for these research questions to be answered, prior assumptions need to be stated and the exact area of research has to be staked out. This is done in the present chapter, which deals with interrogativity as a linguistic function, as well as in chapter 2, which provides an outline of intonation in general and question intonation in particular. On the basis of these introductory chapters, hypotheses are formulated respecting global and local properties of question intonation in Dutch. In chapter 3, these hypotheses are tested with the help of a corpus of production data. Among other things, the results generate the hypothesis that the global shapes of the intonation contours might actually have resulted from specific patterns of accentuation, rather than from independent choices; this part of the analysis is presented in chapter 4. In tandem, the chapters 3 and 4 provide an overview of the intonational properties of interrogativity as observed in the production corpus. Chapter 5 extends the analysis by exploring whether the intonational features of the questions can be partly or wholly attributed to universal biological codes. That is, are these features fully arbitrary, or might they be regarded as (partially) iconic? Also, should the observed properties be analysed as phonological categories or, rather, as part of the process of phonetic implementation? Chapter 6 further elaborates on accentuation, in that it examines the relationship between the questions' accent patterns and underlying focus structures.

[2] That is, the standard variety of that language.

Finally, chapter 7 reviews the results and offers suggestions for further research on the subject of Dutch question intonation.

1.2 The notion 'interrogativity'

On the outset, we should make it clear what, in the present study, is understood by the term 'question'. That is, how can questions be identified, and how do we define the communicative purpose that is being served by them? The issue of what constitutes a question is somewhat blurred by the fact that, in the literature, questions are often approached in terms of their various *categories*. This has spawned a wide range of (overlapping) terms and classifications, and the same category may be referred to by different names: yes-no questions have also been termed 'nexus questions' (cf. Sadock & Zwicky 1985:179), 'inversion questions' (e.g. Geluykens 1988), 'polar questions' (e.g. Huddleston 1994), or 'queries' (e.g. Grice, Benzmüller, Savino & Andreeva 1995). Likewise, wh-questions have also been designated, e.g., 'information questions' (Kiefer 1981:159) or 'variable questions' (Huddleston 1994:417), the latter term indicating that there is a range of potential answers. Discounting such terminological confusion, we still find a wide variety of question types in the literature, ranging from the basic to the truly marginal. Depending on the specific angle of the author, classifications of questions may be based on lexico-syntactic properties, on the intentions of the speaker, on the types of answer the questions elicit, or on the functions they perform. Thus, apart from wh-questions and yes-no questions, mention is made of alternative questions ('….or…..?'), echo questions, tag questions, exam questions, self-addressed questions, guess questions, surprise questions, expository questions, disjunctive questions, existential questions, didactic questions, problem questions, speculative questions, display questions, conducive questions, complementary questions, ditto questions, exhaustive questions, particular questions, requestions, queclaratives, and reclamatory questions (e.g. Kiefer 1981, Wunderlich 1981, Sperber & Wilson 1988:92, Bolinger 1989). Note that this list is far from exhaustive. Further dimensions along which questions have been approached are the cognitive attitudes of the speaker with respect to the expected answer (Kiefer 1981:159), and the social status of speaker and hearer (Athanasiadou 1991).

In much of the literature, an utterance is regarded as a question because of its structure: form is taken to reflect function. That is, it is fairly common for the terms 'interrogative sentence' and 'question' to be equated, since many authors regard questions as exclusively belonging to the syntactic category of interrogative sentences (for an overview, see Huddleston 1994:411). At the same time, these authors are usually forced to acknowledge that not all syntactically interrogative forms function as questions, and, conversely, that utterances functioning as questions may also be realized by non-interrogative forms (e.g. Kiefer 1981:160). Hence, other authors have argued that the terms 'question' and 'interrogative sentence' should be kept strictly apart, considering that they refer to different levels of linguistic analysis (see Hiz 1978:211, Huddleston 1994:411). Under this view, the term 'interrogative sentence' corresponds to syntactic form, the term 'question' to pragmatic function. In the present study, this distinction is of vital interest.

As a first step towards a definition of a question, we will consider questions and interrogativity from the perspective of three of the four standard levels of linguistic analysis, i.e. syntax (§1.2.1), semantics (§1.2.2) and pragmatics (§1.2.3). A consideration from the level of phonetics/phonology will be deferred till chapter 2, which specifically deals with question intonation.

1.2.1 Interrogativity and syntax

Traditionally, sentences have been classified according to their syntactic forms. Indeed, the correlation between grammatical form and conventional communicative use is too obvious to ignore. This has given rise to the classic distinction between three major sentence types which, ideally, are mutually exclusive: the declarative (DECL), the imperative (IMP) and the interrogative (INT) sentence[3]. Each of these is distinct from the others by virtue of a characteristic pairing of formal properties with conventional communicative use. Thus DECL, which is used for conveying information to an addressee, commonly has an overt subject and a finite verb. By contrast, IMP usually lacks a subject while the verb is reduced to its base; IMP is used for giving an addressee orders. INT, finally, is generally used for eliciting a verbal response from the addressee. Formally, it is frequently characterized by inverted word order of subject and finite verb[4]; in addition, there may be a question word or question particle. In this approach, interrogative sentences or, in common parlance, questions constitute a single major syntactic category.

It cannot be doubted that the syntactic classification is insightful, bringing out relationships between formal properties and communicative functions. Also, the fact that the three basic sentence types occur in a majority of the world's languages is indicative of their universal importance (Sadock & Zwicky 1985:181). Yet, as far as questions are concerned it is questionable whether the association between conventional use and syntactic form is strong enough for a purely syntax-based approach. To begin with, though inversion is generally regarded as a characteristic property of interrogative sentences, its presence is by no means compulsory. Certainly, in (1a), a Dutch yes-no question eliciting 'yes' or 'no' for an answer, inversion is indeed obligatory. The same holds for (1b), a wh-question. The question word - which functions as the object here - is fronted, which results in inversion of subject and finite verb.

(1a) *Repareer jij de fiets?*
repair you the bike?

(1b) *Wat repareert hij?*

---

[3] The following generalizations have been taken from Sadock & Zwicky's (1985) survey of 23 languages from various language families and linguistic areas.

[4] The term 'inversion' suggests a specific syntactic operation that changes the order of the subject and the finite verb. However, nowadays the inversion construction is rather seen as a syntactic pattern that is an accidental result of independent syntactic operations, such as the movement of the verb or the topicalization of other constituents.

what repairs he?

By contrast, (1c), though a wh-question, does *not* involve subject-verb inversion. Here, the question word functions as the subject of the sentence. Since in canonical declarative Dutch word order the subject precedes the finite verb, the wh-word automatically fills the initial slot, precluding inversion. In the opinion of Sadock & Zwicky (1985:183), such variable occurrence of inversion in wh-questions weakens the claim that yes-no questions and wh-questions belong to one and the same syntactic category.

(1c) *Wie repareert mijn fiets?*
who repairs my bike?

Conversely, (1d, 1e, 1f) lack inversion although they are meant to function as questions. As regards form, (1d) and (1e) are almost identical in that both have the syntactic properties of a declarative sentence. However, while the former may elicit both agreement and disagreement, the latter, featuring the sentence-final question particle *hè*, only expects confirmation (see also §1.3.2). (1f) is an elliptical question, that is, a conversational question from which subject and predicate have been altogether left out for reasons of economy.

(1d) *Max heeft mijn fiets gerepareerd?*
Max did my bike repair?

(1e) *Max heeft mijn fiets gerepareerd hè?*
Max did my bike repair right?

(1f) *Wanneer?*
When?

Finally, (1g, 1h, 1i) display inversion, in spite of the fact that none of the examples is interrogative. (1g) is a DECL with an adverbial in initial position; (1h) is an exclamative (regarded by some as a separate syntactic category, e.g. for English: Quirk, Greenbaum, Leech & Svartvik 1987:386), and (1i) is an IMP with an overt subject. In Dutch, these syntactic features involve inversion.

(1g) *Morgen repareert hij mijn fiets.*
tomorrow repairs he my bike

(1h) *Had ik mijn fiets maar gerepareerd!*
would that I had repaired my bike!

(1i) *Ga jij eens gauw je fiets repareren!*
go you at once your bike repair!

In sum, it cannot be maintained that the syntactic property of inversion is an essential formal characteristic of interrogativity. On the one hand, inversion is found to occur also in non-questions; on the other, for an utterance to function as a question inversion is not necessary. Rather, by choosing sentence types other than the syntactically conventional speakers seem to be able to add subtle shades of meaning to their questions, commands etc. That there is not a one-to-one correspondence between sentence types and the ways they function has, of course, been convincingly argued by the proponents of Speech Act Theory (Searle 1976). In fact, this has led some theorists to question the claim that such a thing as a limited set of mutually exclusive syntactic sentence types should at all exist (e.g. Sperber & Wilson 1986:247).

If inversion is not an exclusive property of questions, this is equally true of the presence of a wh-word. While in (2) the initial wh-words function as question-words, (3) shows that they are equally appropriate in exclamatory utterances.

(2a) *Wie heeft de Divina Commedia geschreven?*
'Who wrote the Divina Commedia?'

(2b) *Wat heb je liever?*
'What do you prefer?'

(2c) *Hoe moeten we daar komen?*
'How should we get there?'

(3a) *Wie had dat gedacht!*
'Who would have thought that!'

(3b) *Hoe is het mogelijk!*
'How on earth!'

A final counterexample to the idea that questions can be equated with interrogative sentences is the rhetorical question. In most studies dealing with questions it is emphasized that, in spite of its interrogative syntactic form, the rhetorical question should not be regarded as a proper question since it does not require an informative answer (e.g. Kiefer 1980:98; Athanasiadou 1991; Dutch: Droste 1972:128; Haeseryn, Romijn, Geerts, de Rooij & Van den Toorn 1997:1426; but see also Stutterheim 1953:132). Rather, it is a statement that expresses a fairly strong opinion on the part of the speaker.

It seems safe to conclude that, for the identification of questions, lexical and/or syntactic criteria are insufficient. If INT sentences are often characterized by subject-verb inversion and/or by wh-words, so are other sentence types. At the same time, questioning can also be achieved with the help of non-INT sentences.

1.2.2 Interrogativity and semantics

For semanticists, questions have proved difficult to come to terms with and it is not without reason that the semantics of interrogatives has been characterized as an

underdeveloped part of natural language semantics (Groenendijk & Stokhof 1996:1058). Besides, questions may be approached from various semantic angles. For instance, on the assumption that syntactic elements lack intrinsic meaning yet may perform semantic roles, the constituents of an indicative sentence may be analysed as, e.g., 'agent', 'patient' or 'instrument'. With the help of these, the semanticist can establish part of the overall meaning of the sentence. In combination with the meaning of the individual lexical items, he is then able to arrive at the propositional content and to decide on its truth or falsity. However, as many authors have pointed out, questions crucially differ from statements in that they do not contain full propositions. As a consequence, they cannot be assigned truth values (e.g. Hiz 1978:IV, Wunderlich 1981:132, Kiefer 1983:1, Higginbotham 1995:382). In effect, this realisation has hampered the development of a unified semantic account of sentences which also includes questions.

To overcome this problem, it has been proposed that a semantic analysis of questions has to consider the ensemble of question and corresponding answer (see e.g. Hiz 1978:213, Kiefer 1983:1, Higginbotham 1995:369). According to some authors, a question may be seen as representing the set of all propositions which might serve as possible answers (Hamblin 1973), or as true answers (Karttunen 1977, both cited in Groenendijk & Stokhof 1996:1066). According to others, what a question provides is a semantically incomplete proposition which is to be completed by the subsequent answer (for an overview of different approaches, see Wunderlich 1981:136 and Groenendijk & Stokhof 1996). That is, since, by itself, a question cannot represent a certain state of affairs, the proposition has to be constructed in retrospect. Under such views, answers play a vital role in the semantic analysis of questions. However, it should be realised that, in everyday conversation, questions are not necessarily followed by an answer, as the following excerpts from dialogues in English novels illustrate. A listener may choose to ignore a question and respond with an utterance that does not constitute a proper reply at all (4a, 4c, 4d). Alternatively, the speaker may continue the conversation without expecting a reply, even though his previous utterance was interpreted by the listener as a question (4b).

(4a) *"Did all this go on in French?"*
Widmerpool took no notice of this question; which, both Scandinavians knowing some English, seemed to me of interest.
"Örn was more obstinate than Lundquist," said Widmerpool.

(4b) *"But tell me, how do you find Brother Quiggin?"*
I hardly knew what to say. However, Sillery seemed to require no answer. He said: "Brother Quiggin is an able young man, too. We must not forget that."

From: Anthony Powell, *A Dance to the Music of Time* (1951-1955)

(4c) *"[…] How could they have known if they were complete strangers to the district?"*
Petrus chooses not to take this as a question. He puts the pipe away in his pocket, exchanges spade for broom.
From: J. M. Coetzee, *Disgrace* (1999)

(4d) *"I should have been a painter," Walter said. "I showed promise."*
"Did you ever think of getting married?" Matthew said.
"I showed tremendous promise," Walter said, "but my family was indifferent to art.[…]"
From: Muriel Spark, *The Bachelors* (1960)

Evidently, this absence of answers causes problems for a view that the reply is needed to construct the propositional content of a question. That questions do not always receive an answer was also observed *in vivo*, i.e. in a Dutch corpus of spontaneous questions in doctor-patient interactions (Van Heuven, Haan & Pacilly 1998). Not infrequently, questions asked by the doctor were strung together into a sequence of three or more. In his response, the patient usually picked out only one, leaving the others unanswered. In sum, as questions are not always and automatically followed by answers, an approach which crucially includes the replies seems inadequate; in the present study, a question's meaning is taken to be independent of the reply.

If a question does not carry a full proposition, it may yet be related to another semantic notion, that of a presupposition. It is often assumed that questions rest on presuppositions, i.e., on some derivable statement which the speaker holds to be true (e.g. Droste 1972:124, Kiefer 1980:101; for an overview of different approaches to presuppositions, see Groenendijk & Stokhof 1996:1119). Thus, a wh-question is claimed to presuppose a statement in which the wh-element is substituted by an indefinite expression such as *somebody* or *sometime*. For the speaker, this presupposition is the point of departure. The sentence *When is Mary going to France?* is then assumed to contain the presupposition *Mary goes to France some time*, the truth of which is already accepted by the speaker. What the wh-question expresses is that the speaker wishes to add the date of Mary's departure to this presupposition. The answer then causes a new proposition to come into being, for instance *Mary goes to France next week.* A yes-no question may rest on various presuppositions, depending on what is being questioned. For instance, in the interrogative sentence *Does Mary go to France next week?* the scope of what is being questioned may range from the entire presupposition *Mary goes to France next week* (broad focus), to only parts of it (narrow focus). That is, the speaker may only want to know the correct time of Mary's departure, or the correct destination, or whether it is Emma rather than Mary who goes to France next week. In chapter 6, the issue of focus in questions is discussed in more detail.

Presuppositions can be overtly present in the semantics, but they may also manifest themselves contextually. In the latter case, the presupposition is *pragmatic* rather than semantic (Mey 1993:30, Higginbotham 1995:375). In fact, it is suggested by Groenendijk & Stokhof (1996:1122) that semantic theory, rather than aiming at some unified semantic interpretation of questions, should extend its scope

into pragmatics, not only by taking context into account but also the dynamics of information exchange ('dynamic semantics') which, after all, is an essential aspect of interrogativity. Although such a step may blur the dividing line between the two disciplines, it also acknowledges that there is an intimate relationship between meaning and use which should not be ignored.

1.2.3 Interrogativity and pragmatics

Rather than investigating language as an abstract system, pragmatic theory is interested in generalizations underlying the actual *use* of language. That is, sentences are studied from the perspective of the messages they intend to convey in varying contexts. In everyday communication, uttered sentences (i.e., 'utterances') may perform a wide range of so-called speech acts such as warnings, assertions, requests, promises, etc. (cf. Searle 1976)[5]. And, as pragmaticians have been ready to point out, there is certainly no one-to-one correspondence between such functional speech acts and syntactic sentence categories. Rather, as noted above, speakers have considerable freedom in putting the various syntactic sentence types to varying uses, thereby increasing the richness of the communicative resources of the language. For example, it is usually considered more polite to couch a command as a question or a statement, rather than as a straightforward imperative. This opens possibilities for the speaker to express certain attitudes, either towards the listener or towards the message. Given that speakers are free to use syntactic categories in non-categorial ways, it seems more straightforward to regard DECL, IMP and INT merely as categorial peaks along a semantic-pragmatic continuum (Givón 1984). In these prototypical peaks, (syntactic) form maximally coincides with function. However, when syntactic form is less prototypical, the utterance becomes more suitable for performing speech acts other than the canonical. Under such a view, it is possible for the three main syntactic categories to gradually shade into one another (Givón 1984[6], Grønnum 1992).

The insights of Speech Act theory have been crucial in explaining the obvious mismatches between syntactic category and communicative use. In the standard Speech Act account, questions come under the heading of 'requests', that is, the speaker makes a request that the hearer provide him with the relevant and correct information. Yet, it has proved difficult to draw up a universally acceptable Speech Act account capturing all question types (cf. Lyons 1977:755; Mey 1993:151; Sperber & Wilson 1986: 252, Wilson & Sperber 1988:94). For instance, according to some authors questions should be kept distinct from requests for action such as *Can you open the window?*; rather, they should constitute a speech act in their own right (Kiefer 1980:114; Groenendijk & Stokhof 1996:1073). In effect, when it comes to drawing up a unified account of questions pragmatic theory, like semantic theory, is still in some state of flux. It seems

[5] From now on, the terms 'sentence' and 'utterance' will be kept distinct. The former refers to a fairly abstract syntactic/semantic structure, the latter to the actual communicative use that is being made of it.

[6] Givón gives an example of a prototype of the category DECL which gradually shades into a prototype of the category INT: Joe is at home (DECL) > Joe is at home, I think > Joe is at home, right? > Joe is at home, isn't he? > Is Joe at home? (INT).

therefore that, for a viable definition of the subject matter of the present study, additional aspects must be taken into account. This is the subject of the next sections.

1.3 Towards a definition of questions

The present study rests on the assumption that questions constitute a distinct subset of all possible utterances in a language, in that they possess a communicative identity of their own. However, as the preceding sections have indicated, it is not so easy to define that identity. Syntactic and semantic theories do not seem to offer enough substance for a comprehensive definition of questions, at least for the purpose of the present study. In the syntactic framework, the correspondence between syntactic form and questionhood is erratic, making it impossible for questions to be defined on a purely syntactic basis. At most, interrogative syntactic form can contribute towards the prototypicality of a question. At the semantic level, questions have received diverging analyses due to, among other things, the somewhat uneasy relationship between interrogative sentences, truth conditions and propositions. This has caused a number of semanticists to restrict their attention largely to indicative sentences. Besides, in existing accounts of questions an important role is played by the corresponding replies. However, for the present study these are obviously irrelevant: as pointed out earlier, it is not necessary for questions to be followed by a reply. Therefore, it seems more sensible to approach questionhood from the perspective of pragmatics, which concerns itself with the *intended* use of utterances. Accordingly, we view the asking of a question as a unilateral event which rests solely on the speaker's intentions and is separated from a listener's reaction.

Pragmatics offers a further useful starting point in the shape of the element 'information'. When questions are seen as requests for information, they may be considered more or less prototypical as a function of the *type* of information they seek. As Givón has put it, "We don't simply have questions. We have degrees of certainty, degrees of deference, degrees of request and confirmation" (Chisholm 1982:267). Hence, the sort of information sought by a question may be one of the criteria for differentiating between question types. As the present study aims to offer a first systematic description of question intonation in Dutch, it seems sensible to concentrate on the intonation of the most prototypical question types and to leave the more marginal types for future research. This means that, for question types to be eligible for investigation, they should meet certain criteria of prototypicality. Such criteria are formulated in § 1.3.1; § 1.3.2 then briefly describes nine different question types with a view to establishing their degree of prototypicality.

1.3.1 Criteria for prototypicality

As noted above, the pragmatic category 'request for information' can be taken to hold the finer distinction between '(new) information' and 'confirmation (of old information)'. In fact, a request to supply fully new information may sound different from a request to confirm information that is already available to the speaker (Brown, Currie & Kenworthy 1980:189; Gussenhoven 1984a:203; Grice *et al*. 1995; Di Cristo

1998:203). It is important, therefore, to keep information questions and confirmation questions apart. Intuitively, a request for (new) information would appear to solicit a greater transfer of information than a request for confirmation. This would make an information question more prototypical and hence more eligible for the present investigation than a confirmation question. However, a more objective quantification of the notions 'information' and 'confirmation' may be accomplished with the help of the information theory developed by Shannon & Weaver (1948), an approach that is outlined below.

In their classic approach to information as a crucial part of a communication system, Shannon & Weaver roughly define information as the amount of uncertainty in a recipient that is removed by an incoming signal. In their account, information is the logarithm of the number of available choices. For instance, in the case of two equivalent messages, the information equals the logarithm of 2 to the base 2. This unit of information is then called 1 'bit' (binary digit). When we apply this approach to the speech act of questioning, the person asking a question may be taken to correspond to the recipient. When asking a yes-no question the recipient is uncertain between two equally probable answers, *yes* and *no*. Following Shannon & Weaver's proposal, we assume that the speaker's uncertainty can be removed by 1 bit of information, which means that for the speaker each of the two possibilities has a probability of 50% of being eliminated by the reply (i.e., yes-no question = 1 bit). In a wh-question, the speaker's uncertainty is typically greater: to questions beginning with *what, where, when,* etc., there is a whole range of possible answers. That is, this type of question requires more than 1 bit of information: wh-question > 1 bit. By comparison, in a question that seeks confirmation the uncertainty in the recipient is considerably smaller. Since the recipient believes that the answer will confirm a prior assumption of his, the two alternatives for a reply are not equivalent, i.e., not 1 bit (i.e., 50%/50%) but rather, say, 80% for the expected confirmation and 20% for the non-expected disconfirmation, that is, somewhere between 0 bit and 1 bit. Thus, a question for confirmation involves less than 1 bit of information: confirmation-question < 1 bit. Clearly, when information is defined in terms of bits, i.e. in terms of uncertainty that has to be removed in the recipient, it becomes possible to arrive at more precise quantifications of *information* vs. *confirmation*. The assumption that yes-no questions involve exactly 1 information bit, wh-questions more than 1 information bit and confirmation questions less than 1 information bit provides us with objective criteria for selecting question types suitable for the present study.

A further distinction that has been made in the literature is that between primary and secondary questions (e.g. Dutch: Van Es 1932:169). Characteristically, a primary question can be asked independently; there is no need for it to rely on earlier utterances. By contrast, a secondary question always refers back to something that has just been uttered. Thus an echo question, which typically repeats (parts of) a previous utterance, could never be uttered independently. When it comes to selecting the proper question types for this study this distinction seems useful, too. Also, as suggested earlier, questions whose interrogative function is supported by syntactic and/or lexical features can be seen as more prototypical than questions lacking these features.

By considering questions from the speaker's perspective and by differentiating between various question types on the basis of their degree of prototypicality, we are able to specify in more detail what is to be understood by a question in the present study. First, a question is taken to be an *utterance intended to function as a request for information*; it must be distinguished from a request for action. Whether or not the addressee actually *treats* the utterance as a question is irrelevant for our purposes: we regard a question as a unilateral event. Second, with respect to the notion 'information' a distinction is made between questions eliciting (new) *information* (≥ 1 bit) and questions eliciting *confirmation* (< 1 bit). It seems plausible to look upon questions requiring information as more prototypical than questions aiming at confirmation of knowledge that is, in some way or other, already available to the speaker. By the same token, we regard primary questions, which can be asked independently, as more prototypical than secondary questions, which depend on preceding utterances. The same holds for questions which combine their interrogative function with syntactic and/or lexical markers of interrogativity: we consider them more prototypical than questions whose function is not supported by such markers.

Having thus formulated abstract criteria for what is taken to be a (prototypical) question in this study, we will now concretely consider various question types. Taking their formal and functional properties into account, we will then decide which of these types are to be included in the present investigation.

1.3.2 Question types

Although the literature on questions offers a plethora of question types, the following nine are mentioned again and again: yes-no questions, wh-questions, alternative questions, tag questions, declarative questions, echo questions, elliptic questions, rhetorical questions and embedded questions. Apparently, these types are somehow looked upon as belonging to some core, more so than others. Below, they are briefly described; whenever the literature offers specific claims regarding their intonation, these will be mentioned.

Yes-no questions
According to Sadock & Zwicky (1985:179), who compared sentence types across a sample of 23 languages, it is safe to regard the yes-no question as the most basic interrogative type. In Dutch as well as in many other languages, the yes-no question is marked by inversion (e.g. Van Haeringen 1962:290, Haeseryn *et al*. 1997:1427). Seeking agreement or disagreement, it is commonly taken to be unbiased towards y*es* or *no*. (6) gives an example.

(6) *Zitten er forellen in deze rivier?*
'Are there any trout in this river?'

As far as its intonation is concerned, the Dutch yes-no question is, on the main, claimed to have a final rise/rising intonation (e.g. Van Alphen 1914:88; Zwaardemaker & Van Eijk 1928:289; Overdiep & Van Es 1949:87). Furthermore, as

pointed out above a yes-no question typically seeks information rather than confirmation. Finally, a yes-no question can be regarded as a primary question, since it can be asked independently.

Wh-questions

Like the yes-no question, the wh-question is also considered a near-universal question type (Sadock & Zwicky 1985:179). Wh-questions are typically formed with the help of wh-words representing interrogative pro-forms. Referring to some syntactic category, a proform specifies the direction in which the speaker wants to extend the presupposition implied by the question: *who* then refers to a person, *why* to a reason, *how* to a manner, *when* to time, etc. Characteristically, a question-word occurs in initial position, irrespective of its grammatical function. In Dutch statements, the subject slot commonly precedes the slot for the finite verb, as (7a) shows. By the same token, in a wh-question a wh-word functioning as (part of) the subject automatically occupies this initial position (7b).

(7a) statement: *Ellen (SUBJ) vindt de rode bal*
'Ellen (SUBJ) finds the red ball'

(7b) question: *Wie (SUBJ) vindt de rode bal?*
'Who (SUBJ) finds the red ball?'

By contrast, the object in Dutch statements usually occupies a postverbal position. However, it *may* be topicalised to give it additional emphasis. Topicalisation causes the object to be moved into initial position, which then results in inversion (7c); this option is fairly marked, though. In wh-questions, however, utterance-initial position of a wh-word functioning as (part of) an object is the unmarked option (7d).

(7c) statement: *Een páárd (OBJ) heeft hij gekocht*
'A hórse (OBJ) he bought'

(7d) question: *Wat (OBJ) heeft hij gekocht?*
'What (OBJ) did he buy?'

Finally, a wh-word functioning as an adverbial complies with the general rule that adverbials in sentence-initial position go together with inversion, in statements (7e) as well as in questions (7f). Note that, in English statements, a sentence-initial adverbial is not necessarily accompanied by inversion.

(7e) statement: *Gisteren (ADV) vond het meisje de rode bal*
'Yesterday (ADV), the girl found the red ball'

(7f) question: *Wanneer (ADV) vond het meisje de rode bal?*
'When (ADV) did the girl find the red ball?'

Yet, it is also possible for a wh-word to remain *in situ*, i.e. to occur in non-initial position (cf. 7g and 7h):

(7g) question: *Wie (SUBJ) vond wat (OBJ) achter de kastanjeboom?*
'Who (SUBJ) found what (OBJ) behind the chestnut tree?'

(7h) question: *Het meisje (SUBJ) vond wat (OBJ) achter de kastanjeboom?*
'The girl (SUBJ) found what (OBJ) behind the chestnut tree?'

In fact, (7g) is a multiple wh-question in which both the subject and the object are being questioned. Since the initial topic slot is reserved for the subject, it is not possible for the object to be topicalised/fronted. According to Bolinger (1978a:115), such multiple wh-questions are heavily marked. Having two unknowns rather than one, they have to be spoken more slowly and deliberately since their processing require more time.

In (7h), the listener expresses improper understanding or disbelief with respect to part of a preceding utterance. In order to indicate that he expects further clarification he echoes most of the utterance, except for the problematic part which is replaced by a wh-proform; crucially, the syntactic structure remains unchanged.

As regards the intonation of Dutch wh-questions, some authors claim that there is only one type of question intonation, which occurs both in yes-no questions and in wh-questions (e.g. Van Alphen 1914:91; Overdiep & Van Es 1949:490; Den Hertog & Hulshof 1972:139). Others distinguish different intonation patterns. Thus, while some authors claim that wh-questions typically lack a final rise (Guittart 1925:41; Stutterheim 1953:131; Droste 1972:124), others contend that presence or absence of a final rise in a wh-question reflect a speaker's attitude (Zwaardemaker & Van Eijk 1928:289). Since wh-questions can be asked independently of any context, they can be regarded as primary questions. As they crucially elicit new information (i.e. with respect to the presupposition), they are information-questions rather than confirmation-questions.

Alternative questions

In this question type, two or more yes-no questions are coordinated with the help of 'or'; this provides the hearer with two (8a) or more (8b) explicit alternatives. Elements which may be ellipted are parenthesized.

(8a) *Ga je lopen of (ga je) met de fiets?*
'Are you going on foot or (are you going) by bike?'

(8b) *Ga je met de fiets, (ga je) met de bus of (ga je) met de tram?*
'Are you going by bike, (are you going) by bus, or (are you going) by tramway?'

As in wh-questions, the speaker indicates a certain range of possible answers, not by means of a wh-word but by proffering a 'list'. What is implied is, that if the answer to the first part of the question is *no,* the listener has to address the second question,

and so on. However, although an alternative question is a sequence of yes-no questions, the answers *yes* and *no* are not appropriate.

Alternative questions have been reported to have a specific intonation, viz. rising intonation on all non-final elements of the list, opposed to falling intonation on the final element (English: Sadock & Zwicky 1985:179, Bartels 1999:118; Dutch: Van Alphen 1914:92). As a matter of fact, the intonation pattern is quite similar to the so-called 'list intonation' found in statements (cf. Bolinger 1982:11, Bartels 1999:118). Thus, in Dutch as well as in English it is possible for both a statement (8c) and a question (8d) to be realized with a similar sequence of rising and falling intonation (represents rising intonation, represents falling intonation).

(8c) Statement**:**

Engeland is te bereiken met de veerboot, de Hovercraft en het vliegtuig.
'One may get to Britain by ferry, by Hovercraft and by plane.'

(8d) Question:

Gaan we naar Engeland met de veerboot, de Hovercraft, of het vliegtuig?
'Are we going to Britain by ferry, by Hovercraft or by plane?'

Since alternative questions are made up of two or more yes-no questions they are, likewise, of the primary type. Also, they seek information rather than confirmation. What is more, as the 'list' presents different alternatives (each of which corresponds to 1 bit of information), the amount of uncertainty that has to be removed is actually greater than in a single yes-no question. That is, alternative questions require more than 1 bit of information.

Tag-questions

In Dutch, it is possible to turn a statement into a question by appending the particle *hè? ('right?')*. In the literature, such questions are claimed to be requests for confirmation of the information presented in the statement, rather than requests for (new) information (Overdiep & Van Es 1949:491, Van Haeringen 1962:294, Droste 1972, Kirsner & Van Heuven 1996:134). Consider the examples in (9a, 9b).

(9a) *Komt hij morgen ook?*
'Will he be coming too, tomorrow?'

(9b) *Hij komt morgen ook, hè?*
'Tomorrow he'll be coming too, right?'

According to Droste (9a), a yes-no question, can be paraphrased as 'I ask for the truth of his coming tomorrow'. By contrast (9b), a tag question, is to be paraphrased as 'I ask for the confirmation of the truth of his coming tomorrow'. That is, the two sentences would seem to belong to different categories, with (9a) representing an

information question, (9b) a confirmation question. In the tag question, the amount of information needed to remove the uncertainty in the recipient is less than 1 bit. As a tag question first presents the statement that requires confirmation, it can be asked independently of a context; hence, it is of the primary type. As for the intonation of questions ending with the tag *hè?*, Kirsner & Van Heuven (1996:140) have shown that Dutch listeners strongly prefer a realization with a high final boundary tone (H%). Likewise, in a Dutch production corpus *hè?*-questions always featured rising intonation at the end (Beun 1990:45).

Declarative questions

Like the alternative question, the declarative question is sometimes seen as a subtype of yes-no questions (English: Quirk *et al*. 1987:392, Dutch: Haeseryn *et al*. 1997:1428). What distinguishes the declarative question is the absence of formal markers of interrogativity: lexico-syntactically, it is identical to the corresponding declarative sentence, as (10) shows.

(10a) *The bus has departed* (statement)

(10b) *The bus has departed?* (question)

It is widely claimed that, for such utterances to be correctly interpreted as either a statement or a question, the contribution of intonation is crucial (English: e.g. Quirk *et al*. 1987:392; Huddleston 1994:428, but see also Geluykens 1987:493; Dutch: e.g. Van Alphen 1914:91, Van Haeringen 1962:293, Den Hertog & Hulshof 1972:142, Haeseryn *et al.* 1997:1428). To emphasise the functional ambiguity of sentences such as (10), the term 'queclarative' has been coined (Sadock 1974).

It may be mentioned here that, according to some authors, declarative questions are biased in that they are supposed to seek confirmation rather than information, which would render them less prototypical (e.g. English: Quirk *et al.* 1987:393; Huddleston 1994:428; Dutch: Droste 1962:1). However, according to others they must be regarded as information questions, on a par with yes-no questions (Dutch: Van Alphen 1914:91; Den Hertog & Hulshof 1972:142; Haeseryn *et al.* 1997:1428). We propose to adopt the latter position, considering that this view seems to be taken by a majority of the Dutch theorists. As the declarative question does not depend on earlier utterances, it can be seen as being of the primary type.

Echo questions

In this question type, the speaker expresses either surprise/disbelief or improper hearing/understanding of (part of) a preceding utterance (English: e.g. Quirk *et al.* 1987:408; Hockey 1994:99). Therefore, the latter is either repeated almost entirely as in (11a:B), or partially, as in (11b:B); the problematic part is substituted by a wh-proform.

(11a) A. *Ik heb de laatste trein gemist*
'I missed the last train'
B. *Je hebt de laatste trein gemist?!*

'You missed the last train?'

(11b) A. *Ik heb een paard gekocht*
'I bought a horse'
B. *Je hebt een* wat *gekocht?*
'You bought a *what*?'

Usually, there is no inversion, except when an inverted question is echoed rather than a statement, as in (11c:B).

(11c) A. *Waar heb je de paraplu verstopt?*
'Where did you hide the umbrella?'
B. *Waar heb ik de paraplu verstopt? Die heb ik helemaal niet verstopt.*
'Where did I hide the umbrella? I never hid it'

In (11c:B), the addressee expresses surprise at A's question. Instead of answering it he repeats it, indicating that the question was unexpected or inappropriate. Thus, rather than seeking information, echo-questions typically ask for repetition or clarification. In (11a,b), the speaker wants to make sure whether he has correctly heard or understood the previous speaker; at most, this might be taken to reflect a request for confirmation. In (11c), the repeated question merely gives expression to the speaker's surprise; it does not function as a request for information or confirmation. As an echo question always refers back to a preceding utterance, it is intrinsically secondary. In English, echo questions are reported to feature rising intonation (Quirk *et al.* 1987:410).

Elliptic questions

In spontaneous conversation, people frequently omit retrievable elements to avoid repetition. Mostly, an elliptic question is a curtailed wh-question[7]; (12) gives two examples (the B questions).

(12) A. *Ik heb gisteren een tafel gekocht*
'Yesterday I bought a table'
B. *Waar? (heb je een tafel gekocht)*
'Where?' (did you buy a table)'
A. *Ik heb er veel voor moeten betalen*
'It cost me a lot of money'
B. *Hoeveel? (heb je ervoor moeten betalen)*
'How much? (did it cost you)'

[7] There are also elliptic questions without a wh-word, such as 'De buren?' ('The neighbours?'). As a matter of fact, these are often echo questions echoing part of a preceding utterance, for instance 'De hond zit bij de buren' ('The dog is with the neighbours'). Unlike elliptic wh-questions, they do not ask for new information.

According to Van Haeringen (1962:293), Dutch elliptic wh-questions are unlike their full-blown counterparts in that they crucially need rising question intonation. Obviously, elliptic questions are secondary questions; they can only be fully understood when the listener takes a preceding utterance into account. Like full wh-questions, elliptic questions seek new information rather than confirmation.

Rhetorical questions
Lexico-syntactically, rhetorical questions cannot be distinguished from wh-questions or yes-no questions, as (13a,b) illustrate (Quirk *et al.* 1987).

(13a) *Waarom moet het altijd regenen als we willen picknicken?*
'Why must it always rain when we want to have a picnic?'

(13b) *Zullen ze het dan nooit leren?*
'Won't they ever learn?'

Yet, as noted earlier, a rhetorical question is not intended to function as a question; it is a figure of speech typically employed to express some (strong) opinion on the part of the speaker. That is, functionally it has to be regarded as a statement

Embedded questions
Like the rhetorical question, the embedded question may possess formal properties of a wh-question or a yes-no question (wh-word and/or inversion), as (14a,b) show.

(14a) Hij informeerde *wanneer ze naar huis ging.*
'He inquired when she was to go home.'
(14b) Ze vroeg *of Jan naar het ziekenhuis moest.*
'She asked whether John had to go into hospital.'

Crucially, however, an embedded question is subordinated to a main clause stating that a certain speaker is asking a certain question. Since this question is merely reported, not being asked, it cannot be regarded as a functional question (e.g. Paardekooper 1971:30, Droste 1972:128).

1.3.3 Selecting question types for investigation

Considering that the present study involves a first systematic exploration of question intonation in Dutch, we decided to concentrate on prototypical question types. In § 1.3.1, criteria were formulated for assessing the prototypicality of questions; following this, § 1.3.2 reviewed the most commonly occurring question types. Application of the abstract criteria to these question types will enable us to make an appropriate selection.

First, a question whose interrogative function has the support of formal properties (wh-word, inversion) was taken to be more prototypical than a question lacking these formal markers. On the basis of this criterion, wh-questions, yes-no questions, alternative questions, elliptic questions, embedded questions and

rhetorical questions are eligible for investigation. Second, question types seeking (new) information were considered more prototypical than questions seeking confirmation. The amount of information solicited by various question types was objectified with the help of Shannon & Weaver's (1949) information theory. Drawing the line between 'information' and 'confirmation' at 1 bit of information, we intend to select only question types that minimally correspond to 1 bit. This opens the door for wh-questions, yes-no questions, alternative questions, declarative questions, and elliptic questions. Third, so-called primary questions, which may be asked independently of other utterances, were taken to be more prototypical than secondary questions, which hark back to previous utterances. According to this criterion, elliptic questions, echo questions, embedded questions and rhetorical questions have to be ruled out. Table 1.1 sums up the scores of each question type on the three criteria.

Table 1.1. Scores of nine question types on the criteria 'Formal Markings for Interrogativity', 'Primary versus Secondary Question Type', and solicited 'Amount of Information'.[8]

Question type	Formal markings	Primary/Secondary	Amount of information
Wh	*/**	**	>1 bit
Yes-no	*	**	= 1 bit
Declarative		**	= 1 bit
Alternative	*	**	≥ 1 bit
Tag	[*]	**	< 1 bit
Elliptic	*	*	> 1 bit
Echo	(*/**)	*	< 1 bit
Embedded	*/**		
Rhetorical	*/**		

As Table 1.1 shows, wh-questions, yes-no questions and alternative questions have the best aggregate scores on the criteria for prototypicality, followed by elliptic questions. Declarative questions go without any formal markings for interrogativity, which leaves them less prototypical. Tag questions are marked by a final interrogative tag *hè?*, but they explicitly seek confirmation rather than information. Echo questions may or may not be formally marked by a question word, depending on which part of the preceding utterance is being echoed; inversion is rare. However, when functioning as questions (that is, when they do not merely express surprise), they typically seek confirmation, which involves less than 1 bit of

[8] In the second column, '**' indicates the presence of both a question word and inversion, '*' of one of these, '[*]' stands for other formal markings, and '(*/**)' indicates that question word and/or inversion are optional; in the third column, '**' means 'primary question', '*' 'secondary' question; in the last column, the amount of information involved is expressed in bits (see § 1.3.1).

information. As for embedded questions and rhetorical questions, since these do not function as questions they will be left outside further consideration.

If Table 1.1 suggests that wh-questions, yes-no questions and alternative questions should be included in the investigation, the latter type would yet seem to have its drawbacks. Whereas, formally, it can be regarded as a subtype of the yes-no question, it is functionally similar to the wh-question in that it specifies a range of potential answers. In view of this mismatch between form and function the alternative question would appear a less obvious choice. Moreover, there is the suggestion that, from an intonational point of view, the alternative question may be less interesting because it probably features list intonation, which is similar in statements and questions. Considering that the alternative question appears a-typical in more than one respect, it seems proper to exclude it from our investigation. As for the eligibility of the elliptic question, it must be noted that this question type, by definition, constitutes an incomplete sentence. More often than not, it is merely a one-word question, and it seems likely for such reduced questions to present more practical problems for an investigation than do full-blown questions. For that reason, the elliptic question will not be taken into consideration either.

In point of fact, given the purpose of the present study the declarative question seems potentially more interesting, precisely because of its lack of lexico-syntactic markings. Formally identical to a statement, the declarative question is claimed to function as a question solely by virtue of its intonation. This means that it may form an intonational minimal pair with the corresponding statement, permitting clear-cut comparisons. As we aim to investigate question intonation with the help of methods that are maximally controlled, this opportunity should not be wasted. Therefore, it seems sensible to include the declarative question into our investigation, even though this question type is not maximally prototypical.

1.4 Summary

Interrogativity is no doubt a major communicative function. It enables a speaker to elicit information from an addressee and all languages would seem to have a specific subset of utterances giving expression to this function: questions. The present study concentrates on a single aspect of questions, i.e. on their intonation. Specifically, it addresses the question of whether, in Dutch intonation, there are systematic phonetic and/or phonological differences between questions and statements. If such differences are found, it is also investigated whether question intonation varies as a function of question type and of a speaker's sex.

The present chapter provided an introduction to the notions 'interrogativity' and 'question'. Section 1.2 briefly outlined approaches to questions from syntactic, semantic and pragmatic perspectives. Although these explorations yielded useful elements and observations, they did not produce a cut-and-dried definition of questions which could serve as a basis for this study. Therefore, the notion 'question' was further narrowed down in § 1.3.1. Given the advisability of investigating prototypical rather than marginal question types, we formulated three criteria for assessing a question's prototypicality. Next, § 1.3.2 reviewed the commonest question types and § 1.3.3 then

evaluated their degree of prototypicality. Largely on the basis of this evaluation, a subset of question types was selected for investigation.

The observations, decisions and conclusions emerging from this chapter can be summarised as follows:

- Questions are not uniquely determined by form: for an utterance to be a question it need not be marked by inversion and/or a question word. Conversely, there is no need for an utterance marked by inversion and/or a question word to function as a question.
- Questions are defined as requests for information. This implies that, for the speaker, the information elicited by the question is new. When the information is somehow expected by the speaker, we speak of confirmation. As this distinction has been reported to affect intonation, it seems useful to take it into consideration. Crucially, the notion 'question' does not include requests for action, as in *Can you open the window?*.
- A question is seen as a unilateral event and defined as **an utterance which the speaker intends to function as a request for information.**
- Question types combining interrogative function with interrogative form (wh-word, inversion) are taken to be more prototypical than question types that do not. The same position is taken with respect to questions eliciting new information, as opposed to questions merely asking for confirmation, and with respect to primary questions (which can be asked independently), as opposed to secondary questions (which hark back to previous utterances).
- The following subset of question types is singled out for investigation:

  (i) wh-question (WQ);
  (ii) yes-no question (YQ);
  (iii) declarative question (DQ).

  Throughout this study, the term 'Dutch question intonation' refers to these three question types, two of which can be considered as major and prototypical (WQ and YQ); the third (DQ) is included because of its methodological attractiveness.

The next chapter, which is also introductory, looks more closely at intonation, that is, at the functions that can be performed by intonation, at the forms intonation may take, and at the models that have so far been proposed for, notably, Dutch intonation. The chapter further deals with crosslinguistic characteristics of question intonation and their possible origins, as well as with claims regarding question intonation in Dutch. On the basis of these two introductory chapters, hypotheses will be formulated about the intonational properties of the selected set of Dutch question types.

Chapter 2

Introduction and background (2): Questions and intonation

2.1 Introduction and outline

Of a given spoken utterance, intonation forms an integral part. Linguistically, intonation organises the information by indicating, e.g., which parts of the utterance belong together, or which part(s) the speaker means to be of greater or lesser importance[1]. Paralinguistically, intonation may give indications about the speaker's emotion or attitude towards the message or the listener; that is, it may convey politeness, uncertainty, irony, dejection, and so on. Extralinguistically, the overall pitch level of intonation generally indicates whether the speaker is male or female, child or adult. In sum, intonation supplies information with respect to the message proper, to the spirit in which it is expressed, as well as to certain biological characteristics of the speaker. Herein, a spoken utterance crucially differs from its counterpart in print.

The present study concerns itself with intonation in Dutch. More in particular, it focusses on intonational properties of a single communicative function: interrogativity. Presumably all languages have devices to keep utterances expressing interrogativity apart from utterances with the illocutionary force of stating and commanding (Chisholm 1982:278). Thus, interrogativity may be encoded by word order, question words, question particles or question clitics, as well as by intonation. This means that questions constitute a distinct subset of the possible utterances in a language.

Intonation is part of prosody. In order to convey or supplement meaning, intonation often closely cooperates with other prosodic features such as duration and amplitude. Nevertheless, pitch or, in acoustic terms, fundamental frequency (f0) is generally assumed to be the main correlate of intonation (e.g., Bolinger 1986:22, Cruttenden 1986:9, Beckman 1995b:100). Intonation, then, will be informally defined here as the meaningful variation of pitch in the course of an utterance.

Because of its continuously variable nature, intonation has proved elusive to theory[2]. The above working definition pinpoints just two of the manifold theoretical problems intonation theory has (had) to cope with, i.e. meaning and scope. Typically, pitch patterns extend over domains larger than a single segment or word. This has made

[1] Intonation also has an important function in lending cohesion to longer stretches of speech. In this study, however, we concentrate on intonational characteristics of single utterances.

[2] "Intonation […] is one of the most difficult aspects of language sound structure to model." (Beckman 1995a:450)

it hard to decide on appropriate unit(s) of description. Should one make inventories of intonation melodies covering entire utterances ('holistic tunes')? Or, rather, are tunes to be analysed as consisting of smaller component parts? At the same time, establishing the existence of smaller descriptive units seems to require at least some knowledge of which elements can be considered meaningful, and which cannot. This prompts the further question as to whether, by itself, intonation has any meaning. If so, is the relationship between a spoken sentence and its intonation as arbitrary as the relationship between phonemes and words? Or, perhaps, should intonation be regarded as a direct expression of the speaker's mind or attitude?

Answers to such questions may have far-reaching consequences, as they reflect whether intonation is to be seen as belonging to phonetics, phonology, semantics, pragmatics, or any other level of analysis. Obviously, an introductory chapter is not the right place for an extensive discussion of the various aspects of intonation; for this, the reader is referred to handbooks on the subject (e.g. Cruttenden 1994; Ladd 1996). Nonetheless, an investigation of intonational properties of Dutch question intonation necessarily involves issues on which theorists have shown themselves divided. Two such items will be discussed in § 2.1.1, which is concerned with the form of intonation; others are addressed whenever the occasion calls for it. Section 2.1.2 briefly describes two intonation models that have been proposed for Standard Dutch; § 2.1.3 focusses on the relationship between intonation and meaning/function; § 2.1.4 then summarises this first part of the chapter. The second part is more specific in that it concentrates on the intonation of *questions*. After some general remarks in § 2.2, § 2.2.1 looks at features claimed to be characteristic of question intonation, together with theories about their possible origins. Following this, §§ 2.2.2 and 2.2.3 give various examples of the manifestation of high question pitch, in Dutch as well as in a variety of other languages. Section 2.2.4 deals with an additional independent variable in the present study, Sex of Speaker. Finally, in § 2.3, hypotheses are formulated and the chapter concludes with a general summary (§ 2.4).

2.1.1 The study of intonational form

2.1.1.1 Approaches: Production versus perception

The study of intonation is complicated by the circumstance that intonation can be (and in fact has been) approached from various angles and with the help of various methods. For a start, researchers may set out to identify intonational form by exclusively investigating the *production* domain. One obvious thing to do is to perform measurements on the fundamental frequency (f0), on the assumption that pitch is the main acoustic correlate of intonation (cf. § 2.1.2.1). This will yield a wealth of objective phonetic data, which may be plotted as a full-fledged intonation curve. However, it is precisely this abundance of phonetic detail that makes it difficult to identify elements that may be linguistically or communicatively relevant. Also, there are indications that the 'visible' manifestation of pitch does not automatically correspond to the way it is perceived by the listener. Thus, some apparently major pitch event may play a negligible role in perception while, conversely, a seemingly minor phonetic detail may prove indispensible (cf. Cruttenden 1992:525).

To obviate such problems, one may proceed from the opposite end and measure the acoustic correlates of, say, a grammatical feature such as questioning, on the assumption that it is possible for grammatical features to be directly mapped onto the continuously varying f0 curve. To be sure, this may reveal possibly important phonetic properties of questioning. At the same time, however, the method may fail to identify possible invariant elements that also contribute towards the perceived question status of the utterance. That is, the existence of a separate phonological level between the phonetic and the grammatical level is being ruled out in advance (cf. Ladd 1996:23). In the present study it will be assumed, with Ladd, that intonation is like the rest of language in that it has a phonological level of description (Ladd 1996:21).

A third way to deal with production data is to try and identify invariant elements from amongst the phonetic details by careful listening. Other useful acoustic information such as amplitude and duration are then taken into account as well, causing the analysis to have a broader and more realistic basis (but, at the same time, making it harder to determine the contribution of each of the component factors). Obviously, this auditory method is more subjective than the instrumental approach described above; hence, it might give an impression of being less reliable (cf. Lieberman 1965). Also, in the search for clearcut phonological categories, potentially important phonetic features may be missed. However, it is always possible for the postulated categories to be empirically tested with the help of more objective instrumental means (Ladd 1996:13). As the instrumental and the auditory approaches both have their drawbacks, it has been proposed to have the best of both worlds by combining the two (Overdiep & Van Es 1949:85, Grice 1995, Grabe 1998:52)

Another type of access to intonational form is afforded by *perception* judgements on production data that have been modified for the purpose. Again, the idea is that not all phonetic detail carried by the speech signal has perceptual relevance; from part of it, the listener abstracts away. The investigator may attempt to determine which parts are truly relevant by progressively smoothing intonation contours and asking listeners whether the reduced version is still perceptually equal to the original (‘t Hart, Collier & Cohen 1990, see also § 2.1.2.1 below). What remains after this elimination of irrelevant phonetic detail is taken to be indispensable for a correct interpretation. As a further step, listeners have to indicate whether two f0 contours, realized on sentences with different lexical material, are nevertheless perceptually equivalent, that is, whether they are similar enough to be judged as instantiating the same contour type. This method makes it possible to uncover invariant formal units as well as their possible combinations. Note, however, that it accesses the phonetic surface; it does not seek to identify possible abstract underlying forms and structures.

To sum up, an acoustic production approach yields an abundance of objective phonetic data. However, in order to identify the linguistically relevant parts the investigator has to carefully ‘prune’ the data, thereby relying largely on his own intuitions. In an auditory production approach, the signal is filtered by the investigator who, also with the help of his intuitions, attempts to identify invariant elements. Ideally, the two approaches should be combined and the units that have been identified should then be validated with the help of perception experiments. This, however, has not always proved feasible. By comparison, an approach based on multiple perception

judgements by native speakers would appear more objective. On the listeners' directions, the set of production data is gradually stripped of phonetic detail and what is left is taken to be the relevant part of the original input.

As it is, all of these methods have been employed in the descriptions of Dutch intonation; more on this subject will be said in § 2.2.3 below.

2.1.1.2 Units of description

As pointed out earlier, identifying intonational primitives has proved a major challenge for intonation theory. During the past century, various proposals have been made which will not be discussed here, as they are not immediately relevant to the subject of the present study (for surveys, see Cruttenden 1994; Ladd 1996). What is important here, though, is whether intonation should be seen as some global device covering, say, an entire utterance, or whether it builds up step-by-step, linking up smaller units.

Broadly speaking, there has been a tendency in intonation theory to move from full contours towards increasingly smaller units of description. Especially in the earlier literature, intonation was often equated with sentence intonation, that is, the grammatical meaning of a sentence was believed to shift as a function of its overall 'melody'. These holistic analyses had their problems, though. For a start, in instantiations of certain melodies, the lexical material alone tended to be conducive to the 'correct' interpretation, or a fairly elaborate context. Thus, the danger of confusing lexico-syntactic content with intonation was considerable (cf. Couper-Kuhlen & Selting 1996:20). Also, partial intonational similarities between different tunes often went unnoticed, which meant that possible generalizations were missed. From a psycholinguistic point of view, a holistic position would also seem problematic because, in the worst case, it would predict the existence of a long list of all possible melodies in the language, comprising all formal and pragmatic variants in combination with all paralinguistic gradations. This seems unlikely, as such a list would impose a heavy load on the processing of sentences (cf. Cutler 1977).

A number of investigators preferred to analyse the intonation of an utterance as being made up of smaller component parts. Some postulated four local pitch **levels**, Low, Mid, High and Overhigh, and sentence intonation was seen as a composite of these (American Structuralist approach of the 1940s and 1950s, cf. Ladd 1996:60). By others, sentence intonation was analysed as consisting of a sequence of local, distinctive pitch **movements**, i.e. of various types of rises and falls. The distinctiveness of these movements had been previously established by means of visual, auditory and/or perceptual methods (e.g. the 'British School' approach, from the 1920s onwards, cf. Bolinger 1989:3; Grice 1995:7).

Also, it has been proposed to combine a 'local' approach with a 'global' component. The assumption was that a phrase or utterance may have overall acoustic characteristics (notably the global direction of the f0 course), upon which the speaker then superimposes local falls and rises (the so-called 'overlay' or 'superposition' models, e.g. 't Hart & Collier 1975, Fujisaki 1983, Gårding 1983, Thorsen 1980/Grønnum 1992, 1995). By some, the overall course of f0 was claimed to have a linguistic relevance of its own, notably in questions (see § 2.2.2.2 below).

From the late 1970's onwards, Bruce (1977), Pierrehumbert (1980) and others further reduced the number of intonational primitives by positing only two contrasting units, a high tone (H) and a low tone (L). These tones were assumed to specify pitch *targets*, that is, at a certain point in time pitch has to reach a certain level[3]. While, phonologically, this level is thus either High or Low, its actual phonetic value arises from phonetic implementation rules which may produce 'allotones' out of H and L, as a function of their tonal context or of certain prominence factors (Pierrehumbert 1980; see also chapter 5, § 5.3.5). Between the high and the low tones, f0 is simply interpolated, thereby producing what others may describe as a rise or a fall. That is, in this framework pitch movements are merely derivative, they are not primitives in themselves. It is possible for the two phonological tones H and L to be combined into pitch accents, e.g. L+H*, with the asterisked tone associating with an accented syllable. Moreover, it is suggested that the target tone is not only variable as regards its actual f0 (i.e. vertical variation, or 'scaling'), but also with respect to the point where it aligns with the accented syllable (i.e. horizontal variation, or 'alignment'). That is, a phonologically motivated pitch peak (H*) may occur prior to, within or after the accented syllable. The factors governing alignment are currently objects of research (cf. D'Imperio 1997; Arvaniti, Ladd & Mennen 1998, Wichmann, House & Rietveld 2000). In all, the autosegmental approach has been a powerful influence in intonation theory; for more details, the reader is referred to Ladd (1996).

2.1.2 Models of Standard Dutch Intonation

It is not surprising that studies of Dutch intonation show reflexes of the general developments in intonation theory. Although, early on, it was claimed that phonological oppositions could exist between entire sentences, rather than within paradigms of local intonation units (cf. Van Wijk 1939:142), there have been shifts from a holistic towards a more atomistic approach, as well as from a concrete approach towards a more abstract (autosegmental) approach. The latter shift, presumably, was bound up with the emergence of phonology, reflecting a growing need in linguistic research to isolate parts that are used to differentiate meanings from parts that are not (cf. Van Wijk 1939).

In 1925, Guittart made what is presumably the earliest attempt to describe Dutch intonation. Using the auditory method (cf. § 2.1.1.1) and distinguishing five sentence types (declaratives, imperatives, exclamatives, interrogatives and optatives), he indicates for each of these types whether the nuclear pitch movement (the most important tonal part) is generally falling, rising, rising-falling-rising or weakly rising. That is, like many of the earlier investigators of intonation, Guittart looked for direct correspondences between intonation patterns and grammatical sentence types. At times, however, the intonation patterns he posits would seem to be influenced by the lexical content of the utterances in question (cf. Couper-Kuhlen & Selting 1996:20). Guittart's auditory

[3] Tones were also called 'tone segments' or 'autosegments', reflecting that they were viewed as autonomous segments occupying a representational tier of their own; hence the term 'autosegmental theory' (cf. Goldsmith 1976).

approach was denounced by Van Es (1932:123), who urged that intonation can be fruitfully studied only by means of experimental methods. As an example, Van Es held up a pilot experiment of his own, in which a neutral question had been recorded together with four 'emotional' variants (expressing, e.g., surprise and fear). Crucially, the segmental part had been kept constant. By measuring f0 values and durations at certain points, Van Es was able to pinpoint local and global prosodic differences between these five variants. What Van Es's fairly severe criticism of Guittart's auditory approach illustrates is that, for a modern investigator of intonation, it was now considered unwise to go on relying exclusively on traditional methods.

Below, the early concern with intonation as a prime indicator of grammatical sentence type in Dutch will be discussed in more detail (§ 2.1.3). First, however, we will concentrate on the two models of Dutch intonation that are currently regarded as leading, the so-called IPO Grammar of Dutch Intonation ('t Hart, Collier & Cohen 1990) and the autosegmental model as expounded in, e.g., Gussenhoven 1984a, 1988, 1991, Gussenhoven & Rietveld 1992, Van Den Berg, Gussenhoven & Rietveld 1992).

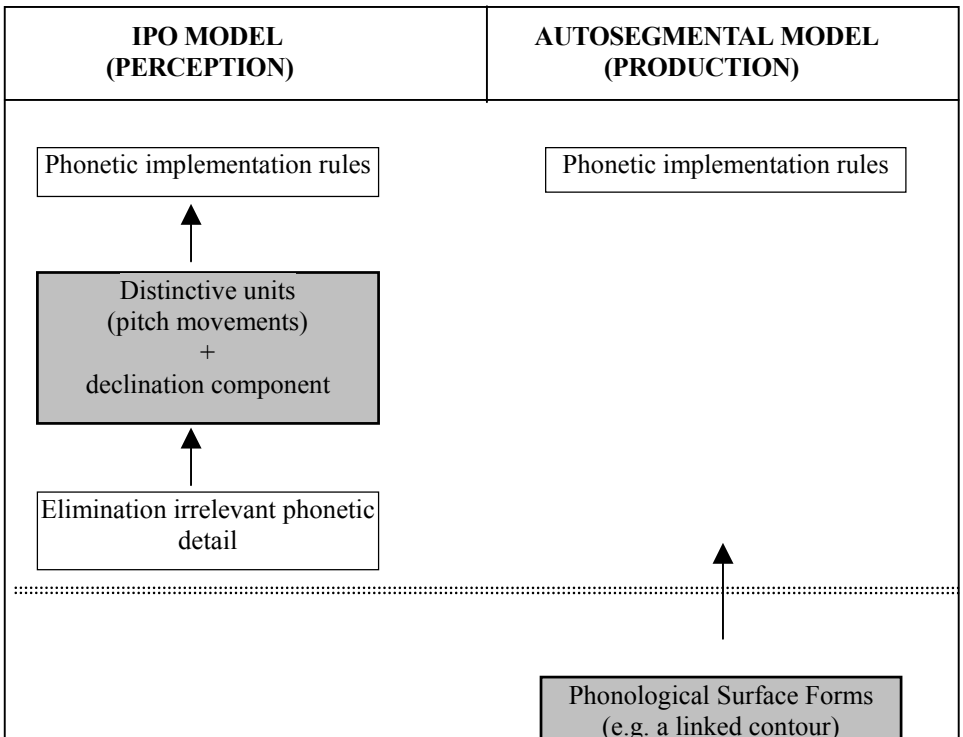
2.1.2.1 The IPO Grammar of Dutch Intonation

Roughly between the 1960s and the late 1980s, researchers at the Eindhoven Institute for Perception Research (IPO) developed a comprehensive description of Dutch intonation, which was firmly rooted in perception ('t Hart *et al.* 1990). Using the method described in § 2.1.1.1 above, the authors isolated ten basic movements, five rises (indicated by 1, 2, 3, 4, 5) and five falls (indicated by A, B, C, D, E). Each of these has a (+) or (–) value for a number of phonetic features, such as occurring [+early] vs. [–early] in the syllable, or encompassing the full pitch range vs. part of the pitch range. The ten movements can be put to various uses. Thus, while some of the falls and rises create accents, others may connect movements with one another, or take high pitch back to the low base line. The half-sized rise '5' can act as a boost to the full-sized rises '1' and '4', and the rise '2' produces a high boundary tone (cf. 't Hart & Collier 1975; Cruttenden 1992; 't Hart *et al.* 1990). The movements may be combined to form configurations such as '1&A' (an accentual rise immediately followed by an accentual fall). Configurations, in turn, may be strung together to form full contours consisting of an optional Prefix, an obligatory Root and an optional Suffix ('t Hart *et al.* 1990:79); these contours are the largest descriptive units. As not all combinations of movements and configurations are permitted, the grammar contains a number of syntagmatic constraints. Finally, it is assumed that movements or configurations are superimposed on an overall downward f0 trend, the 'declination' component (cf. § 2.1.1.2 above and § 2.2.2.2 below). This declination is taken to manifest itself in the topline (which connects the f0 maxima of the movements and is transcribed 'Ø'), as well as in the baseline (which connects the f0 minima of the movements and is transcribed 'O').

The IPO descriptive units are assumed to represent only the planned (and hence communicatively relevant) changes in pitch[4]. Consequently, they are regarded as

[4] An important principle underlying the body of IPO research was that "only those f0 changes would be regarded as possible candidates for a descriptive model of pitch for which a link could be established with commands to the vocal-cord mechanism, which as such are under the speaker's control" (p. 186).

the phonological units of Standard Dutch, or, as one of the researchers puts it, "[…] the 'phonetics of pitch' is at the same time the 'phonology of intonation' " (Collier 1989:256). That is, the IPO model consists of a single level of representation which is at once phonological and phonetic; there is no abstract underlying level. Below, the left half of Figure 2.1 schematises the IPO approach. Starting from the phonetic surface, involuntary and hence irrelevant pitch is removed with the help of listeners' judgements. What remains, after this procedure, are the pitch movements that are taken to be communicatively relevant, together with a global declination component (shaded box). The approach can be characterised as 'bottom-up' since it takes the phonetic surface for a starting point. The invariant formal units then constitute the end point; they are supplemented with a set of phonetic implementation rules (non-shaded box).



surface level

underlying level

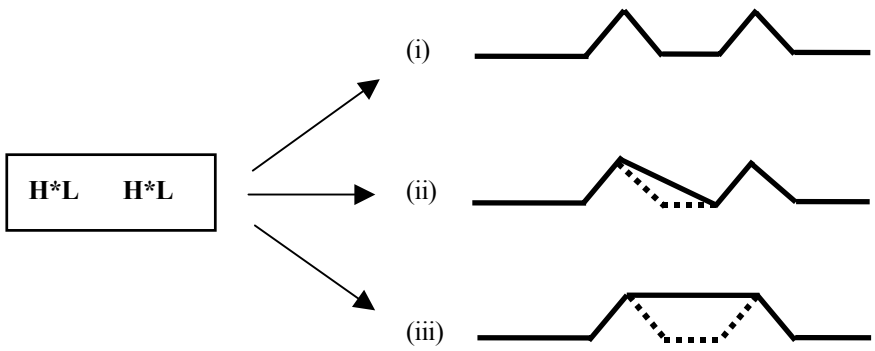
Figure 2.1. Schematic representation of the two leading approaches to Dutch intonation, the IPO model (‘t Hart, Collier and Cohen) and the autosegmental model (Gussenhoven and colleagues).

2.1.2.2 The autosegmental model

Unlike the IPO model, the autosegmental model of Gussenhoven posits a separate underlying phonological level which is subdivided into two parts (cf. the right half of Figure 2.1). The lower phonological level contains the abstract basic tones H and L (‘tone segments’). These may combine to form the bi- and tritonal ‘tone words’ H*L, L*H and H*LH, each of which may be selected by a speaker to realise a pitch accent (lower shaded box). Crucially, however, it is possible for these tones to be modified by a number of explicitly formulated operations (‘rules’ and ‘modifications’, Gussenhoven 1988:98). The output forms of these operations then appear at the second phonological level as surface phonological forms (upper shaded box). In other words, the underlying level of representation comprises basic phonological categories as well as derived phonological surface forms. Phonetically, the tone words are to be seen as combinations of individual high or low tonal targets (see section 2.1.1.2 above); from one tonal target to the next, f0 is linearly interpolated (for further details, see Gussenhoven and Rietveld

1992). The actual f0 value of the tonal targets is determined by a set of phonetic implementation rules; this component is what both models have in common. The idea behind the two-level phonological approach is that the basic tone word(s) can be moulded to (subtly) different communicative ends with the help of certain intonational 'morphemes', such as downstep. Also, during the process of speaking, the intonational phonological units may be subject to local adjustments in the same way as their segmental counterparts[5].

Importantly, the two phonological levels enable Gussenhoven to express greater generalizations. Thus, it is possible for seemingly dissimilar phonological surface forms to be reduced to one and the same underlying form. This is appealing, as many surface forms are intuitively felt to be variants of a single basic form, rather than different phonological forms requiring different transcriptions (as may be the case in the IPO model, as well as in the autosegmental model of American English of Pierrehumbert and colleagues). To give an example, Gussenhoven argues that a sequence of the two tone words H*L H*L at the lower phonological level may have three different variants at the upper phonological level, two of which result from phonological processes. This is shown in Figure 2.2.



[5] According to Gussenhoven, the southern Dutch dialect of Venlo presents further support for two phonological levels. The dialect in question has lexical tone, that is, tones are not merely employed for intonational purposes but also for making lexical and grammatical distinctions, such as between singular and plural. Especially at utterance ends, lexical tones and intonational tones can be seen to interact. As a result, the lexical tones may show up in radically different shapes, depending on the intonational units they are being combined with. For an adequate description of some of the Venlo phenomena, one needs two separate phonological levels (Gussenhoven & Van Der Vliet 1999:133).

Figure 2.2. Tone linking (after Gussenhoven 1991:145). A basic sequence of H*L H*L at the underlying phonological level (left box) may appear in three different shapes at the surface phonological level. Dotted lines reflect deletion.

In Figure 2.2, (ii) differs from basic (i) in that the boundary dividing the two accent domains has been removed, enabling the L tone of the first pitch accent to shift rightwards to some point just before the rise of the second pitch accent. In effect, the two domains have merged into a single one. Next, (iii) differs from (ii) in that the morpheme Tone Linking has been added, deleting the L tone of the first accent (following the merger of the two domains). Such an analysis is more economical in that it does not require three different transcriptions. Also, it is more insightful in that it does justice to the intuition that the contours (i, ii, iii), broadly, belong together, even if their formal differences correspond with (subtle) pragmatic differences.

By and large, Gussenhoven's model is an exponent of the autosegmental framework. This means, among other things, that intonation contours are seen as resulting from sequences of abstract tones associating with accented syllables, the phonetic realisation of which is to some extent variable, both in terms of scaling (actual f0 value) and of alignment (actual place of association with accented syllable)[6]. It is worth noting, however, that, unlike many other autosegmental descriptions, Gussenhoven's model refers to meaning; more will be said about this aspect in the next section. It may further be added that, recently, the phonological contrasts proposed in Gussenhoven (1984a, 1988, 1991) have been subject to some changes and additions on the occasion of the development of the ToDI system (Transcription of Dutch Intonation; Gussenhoven 2002b, to appear). Thus ToDI, which means to provide an exhaustive description of the phonological contrasts of Dutch, includes the contrast between L*H H% and H* H% (corresponding to L* H-H% vs. H* H-H% in Pierrehumbert 1980), established in Gussenhoven & Rietveld (1997). Furthermore, ToDI transcribes the non-final occurrence of pitch accent H*LH as H*+L (to describe a steep fall followed by a slow rise to another pitch accent), and modification [DELAY] is now represented by the prefix L* (cf. Gussenhoven & Rietveld 1992). Importantly, ToDI identifies the contrasts found in the phonological surface component (cf. Figure 2.1). These contrasts may be products of earlier phonological or morphological processes that have affected the more abstract underlying units in the lower box of Figure 2.1. In ToDI, the effects of such abstract tonal modifications are expressed directly in the tonal representation. Thus, for instance, the high plateau of the 'flat hat' in contour (iii) in Figure 2.2 above is not transcribed as the abstract, underlying H*L, but as H*, which reflects sustained, 'spreading' high pitch. More details about ToDI are given in chapter 4, § 4.2.1.

2.1.3 Intonation and meaning

Time and again, investigators have commented upon the difficulty of pinning down the *meaning* of intonation. Indeed, it has been claimed that intonational meaning is so dependent on context that pairing specific contours with specific meanings seems an almost impossible undertaking (Cutler 1977:113). Others contend that "there is little or

[6] For a comprehensive overview of autosegmental and post-autosegmental theory, see Grice 1995:57.

no constancy between intonational form and meaning: in the worst case, the same tune 'means' something different with each different lexico-syntactic carrier" (Couper-Kuhlen & Selting 1996:20). To be sure, a listener's interpretation is strongly guided by both context and lexical content. It is also true that many of the meanings hitherto attributed to intonational tunes or elements have been based on intuitions rather than on strictly controlled experiments. Nevertheless, it has proved possible to systematically vary the intonation of lexically or contextually identical material and to conclude that certain formal elements may have some constant (pragmatic) effect on a listener's interpretation (Dutch: e.g. Caspers 1996, 1997, 1998; Grabe, Gussenhoven, Haan, Marsi & Post 1997). Evidently, it is possible for elements of intonational meaning to be uncovered by rigorously excluding lexical and contextual variation.

It is generally felt that linguistic meanings of intonation should be kept distinct from paralinguistic meanings, the latter being phonetic/gradient rather than phonological/categorical and comprising emotions and attitudes such as anger, surprise, mistrust, etc[7]. According to Pierrehumbert & Hirschberg (1990:284) and Ladd (1996:39), paralinguistic meaning cannot be interpreted from the contour in a direct fashion. Rather, the listener constructs it on the basis of the linguistic meaning of the contour together with the overall context. Linguistic meaning, by contrast, is taken to be directly interpretable from the formal categories making up the f0 contour. These categories would seem to have fairly basic meanings having to do with, broadly, the informational import of the utterance, i.e., how the hearer is meant to process and interpret the speaker's utterance (with respect to context and a mutually shared background). In Pierrehumbert & Hirschberg's model, pitch accents, phrase accents and boundary tones are assumed to have scope over different domains of interpretation and to convey specific types of abstract linguistic meaning. To give an example, "The L* accent marks items that S[peaker] intends to be salient but not to form part of what S is predicating in the utterance" (p. 291). Building on these assumptions, Pierrehumbert & Hirschberg propose a compositional approach to intonational meaning, which is primarily based on Pierrehumbert's inventory of phonological categories (e.g. Pierrehumbert 1980; Beckman & Pierrehumbert 1986). The idea is that the successive chunks of meaning guide the listener towards a correct interpretation of the complete utterance. Thus, both at the formal and at the functional level intonation is viewed as sequential: it develops step-by-step. This sequentiality is one of the key notions of autosegmental theory.

As for the possible meanings of formal intonational elements in Dutch, the IPO grammar has explicitly left meaning or function outside consideration[8], among other things because of sundry methodological problems ('t Hart & Collier 1975:254; Collier

[7] Emotions such as fear and sadness can be defined as primary reactions to some event which do not involve an active choice by the speaker. By contrast, attitudes such as irony or scorn. would seem to have a cognitive component: the speaker more or less consciously adopts a certain attitude towards the message or the hearer. Still, it is not always easy to draw the line between the two. 'Surprise', for instance, would seem to fit both categories.

[8] Proposition 9 of *A Theory of Intonation* says: "Intonation features have no intrinsic meaning" ('t Hart, Collier and Cohen 1990:110).

1989:257; 't Hart *et al.* 1990:185). According to the authors, "the assessment of the formal properties of intonation takes logical precedence over the study of its linguistic and expressive use" ('t Hart *et al.* 1990:5). They argue that intonation differs from the segmental part of speech in that the relationships between form and meaning are many-to-many, and there seems to be no such thing as a lexicon which, in segmental studies, has proved extremely helpful in separating out the various categories. Things are further complicated by the fact that differences in meaning as conveyed by intonation may be quite subtle. Still, the authors are aware that establishing formal intonational categories without having some idea as to whether their 'meaning' somehow contrasts with the 'meaning' of others may hold a risk of over- or underdifferentiation. It may be noted, however, that explicit form-meaning correspondences for intonational units posited by the IPO have been offered by Keijsper (1984)[9].

Unlike the IPO description, the autosegmental description of Dutch intonation does concern itself with meaning, be it in a fairly abstract way. The three tone words in Gussenhoven's account carry the general meanings ADDITION, SELECTION and TESTING (Gussenhoven 1984a, 1988), which refer to the status of the accented items with respect to the common background of speaker and hearer. For example, the tone word H*L is taken to introduce an item to the shared background; this makes it suitable for statements. Similarly, the rules and modifications ('intonational morphemes') pass on some constant element of meaning to the tone word they modify. Nevertheless, as said earlier it is not always easy to experimentally verify the proposed meanings of these tone words and morphemes.

In sum, where the developers of the IPO grammar have explicitly declined to take meaning into account, Gussenhoven's description crucially relied on meaning, however abstract. At present, both positions would seem to have somewhat relaxed. Thus, it has been acknowledged that the IPO grammar of intonation overgenerates and needs to be constrained by the demands of communication. Therefore, a shift of attention is needed from the description of phonetic form to the study of communicative function (Collier 1993:67). Gussenhoven (1999) on his part indicates that "Intonation combines a high degree of analytic indeterminacy in the phonology and morphology with a similar high degree of analytic indeterminacy in its semantics." In short, while establishing intonational form is not an easy task, intonational meaning would seem to offer precious little to hold on to.

In the present study, it seems sensible to subscribe to what Ladd (1996:39) has called the 'Linguist's Theory of Intonational Meaning': "The central idea of this view is that *the elements of intonation have meaning.* These meanings are very general, but they are part of a system with a rich interpretive pragmatics, which gives rise to very specific and often quite vivid nuances in specific contexts." To a large extent, linguistic meanings as expressed by intonation would seem to be discoursal, i.e. they broadly relate to ways in which an utterance may affect the body of mutual knowledge shared by speaker and hearer. In fact, the production of a statement presupposes some cognitive imbalance between the interlocutors, and by uttering the statement, the speaker aims to update the hearer's current state of knowledge. In *questions*, it is the reverse: the speaker's background is out of step with that of the hearer. Unlike a statement, a

[9] For experimental verification of Keijsper's proposals, see Caspers 1996, 1997, 1999.

question seeks to resolve some presupposition on the part of the *speaker*'s, and it seems plausible for this important discoursal difference to be reflected in the intonation. Thus, interrogativity might be seen as one of the possible 'meanings' signalled by means of intonation.

2.1.4 Summary

The first part of the present chapter was concerned with intonation. After a general introduction (§ 2.1), two issues were briefly discussed. First, an outline was given of the methodological problems posed by the extremely variable outward forms of pitch contours (§ 2.1.1.1). Next, § 2.1.1.2 dealt with the various units of description that have so far been proposed. It was seen that, in the course of time, there has been a tendency for these units to become smaller and more abstract. Competing views were discussed, notably the view that intonation consists of rises and falls superimposed on a global downward f0 trend, as opposed to the view that an intonation contour arises from a strictly linear sequence of individual high and low tones, each of which specifies a phonetic target. In the latter outlook, rises and falls are merely by-products of interpolating between such tonal targets. The two current models of Dutch intonation can be broadly seen as exponents of these two views (§§ 2.1.2.1-2).

Thus, the IPO Grammar of Dutch Intonation largely rests on perception judgements of intonation contours that have been stylized with a view to removing all supposedly irrelevant phonetic detail. What remains of the speech signal is then assumed to be planned by the speaker and, hence, relevant for the listener. This method has yielded an inventory of ten formal units (pitch movements) and their possible combinations. The movements are claimed to be superimposed on a globally downward f0 trend, irrespective of utterance type. In all, much effort has been directed towards establishing an objective and reliable method for identifying the basic units of (Dutch) intonation. The autosegmental model (Gussenhoven and colleagues) was found to be more abstract. Its three basic 'tone words' (which are claimed to have certain discoursal meanings) may be involved in a number of morphological operations. While this requires two different phonological levels, it opens the possibility to derive surface forms which, intuitively, are felt to be related to one of the basic, underlying forms.

Section 2.1.3, finally, discussed the somewhat uneasy relationship between intonation and meaning. Although, in some of the earlier analyses, linguistic and paralinguistic meaning components were lumped together, the distinction between the two is now firmly established. In the autosegmental framework, it is generally accepted that intonational categories such as pitch accents and boundary tones have meanings of their own, be it that these are fairly basic. Such meanings broadly relate to the way in which the categories are meant to be interpreted vis-à-vis the mutual beliefs of speaker and hearer ('discoursal meaning'). As for the models of Dutch intonation, while the IPO Grammar has typically kept away from meaning, the autosegmental model has proposed fairly abstract discoursal meanings for each of the three postulated tone words.

2.2 Intonation in questions

One must view with suspicion any claim that a language has absolutely no pitch-marking of questions.

D. Bolinger, 'Intonation across languages'.
In: J. Greenberg (ed.) *Language universals (1978:501)*

If, so far, intonation theory would seem to have treated meaning and function fairly circumspectly, question intonation proves the exception. Since, in any language, interrogativity is a clear-cut functional category accompanied by distinct morphosyntactic and/or prosodic features questions have, of old, appealed to linguists. As regards the intonation of questions investigators have, in the main, concerned themselves with two related issues. First, can interrogative sentences be distinguished from other sentence types by means of intonation, and second, if there is a distinct question intonation, what are its characteristics? The former issue is briefly addressed in the present section; the latter is discussed in § 2.2.2 below.

With respect to the alleged distinctive power of question intonation, opinions seem to range between extremes. On the one hand, it has been claimed that intonation is crucial for keeping grammatical sentence types apart (e.g., English: Halliday 1967a:10). Others, in contrast, have maintained that "any intonation that can occur with a statement, a command, or an exclamation can also occur with a question" (Bolinger 1989:98), and that "questions may be found with all intonation curves" (Pike 1945/1972:163; also: Geluykens 1988:468; Cruttenden 1994:59, 96). Note, however, that there are also intermediate positions.

As regards Dutch (question) intonation, De Groot (1943:31) and Overdiep & Van Es (1949:84) take the stricter view that, for the expression of a sentence's modality, intonation is decisive. Less rigorous positions are taken by authors who claim that only in declarative questions, or in elliptic questions such as *What else?,* is intonation truly indispensable and hence linguistic (Van Alphen 1914:91, Van Wijk 1939:143, Van Den Berg 1958:124, Droste 1961:12, Van Haeringen 1962:289). The view that there is no such thing as a specific intonation for questions and that it is impossible to distinguish questions and statements on the basis of their intonation, is expressed by, e.g., Stutterheim (1953:135), Paardekooper (1971:31) and 't Hart (Hirst & Di Cristo 1998:103). Thus, for instance, 't Hart notes that final rises, often viewed as the defining property of question intonation, occurred in no more than half of the 100 questions of the IPO corpus, whereas non-questions frequently featured final rises signalling continuation.

2.2.1 (Universal) properties

'Erotesis [interrogativity], *if it be pure, raiseth the common tone or tenour of the voice in the last word;[…] but if it begin with a word interrogative […] it falleth as a period'.*

Butler (1634), as cited in Cruttenden 1981:78

One of the first linguists to occupy himself with the intonational properties of questions in a wide variety of languages is Kretschmer (1938). He states that, in general, question intonation can be characterized as a final rising tone, a feature reported for many languages across the world. As regards the origin of this widespread property, he firmly disagrees with Wegener's earlier claim (1885, cited in Kretschmer) that the rising tone should reflect (paralinguistic) surprise. For Kretschmer, this assumption leaves unexplained why both wh-questions and disjunctive questions (i.e. double questions linked by 'or'; 'alternative questions') lack the final rising intonation, though both are questions. Instead, he argues, a yes/no question such as *Did he come home?* should be interpreted as the disjunctive *Did he come home or did he not come home?*[10]. Given that the first part of this double question ends in a rise, the second in a fall, a yes/no question should simply be seen as a truncated version. That is, only the first part is realized; the second part with falling intonation is ellipted causing the rising final intonation of the first part to become a characteristic cue to interrogativity. With a similar argument, Kretschmer explains the intonation of wh-questions. In full-blown versions (*What did you see?*, *Why did you go?*), the initial question word has rising intonation[11] but following that, pitch falls. In effect, this question type has the overall falling intonation of a statement. In ellipted versions (*What?, Why?*), however, the question word retains its rising intonation; the falling intonation of the second part is lost as a result of the ellipsis.

Kretschmer finds evidence for his view in the circumstance that, in some languages, questioning is (still) solely expressed by disjunctivity[12]. Also, in many languages question particles can be shown to be actually disjunctive particles, broadly meaning 'either'[13]. Besides, yes-no questions are often followed by a disjunctive phrase meaning…, *or (not)?*; this must, likewise, be interpreted as a residue of full disjunctivity. In sum: rising question pitch in yes-no questions is the product of ellipsis, and Kretschmer's account neatly explains why rising pitch can also be found in ellipted wh-questions, whereas full-blown wh-questions and disjunctive, 'alternative' questions go without it. Kretschmer then goes on to say that rising final pitch in (yes-no) questions is found in a great many languages throughout the world. It is not immediately obvious, however, why the initial part of alternative questions should have (had) rising intonation in the first place, or how the rising pitch of initial wh-words should be accounted for. Finally, with respect to tone languages Kretschmer expresses the view that the use of f0 for making lexical and/or grammatical distinctions[14] is inherently incompatible with the

[10] According to Kretschmer, disjunctive questions have evolved from disjunctive statements. Inversion is believed to have developed much later, around the Middle Ages (Kretschmer 1938:39; see also van Haeringen 1962:295).

[11] Kretschmer established the rising intonation in the question word with the help of an oscillograph.

[12] Cf. Li, in Chisholm 1982:255: "Well, actually the most general type that's found in languages in terms of yes/no questions is the alternative question." (Interrogativity, Plenary session).

[13] The English *whether* is a case in point

[14] In Thai, for instance, the sequence /kha:/ can mean *to trade, a kind of grass, galangale, to kill*, or *a leg*, depending on whether it is produced on a high, a mid, a low, a falling or a rising tone (Luksaneeyanawin 1998).

final rising intonation of questions. However, to this claim many counterexamples have since been produced.

Hermann (1942), who gives a comprehensive overview of question intonation in 175 related and unrelated languages across five continents, challenges Kretschmer's view that question intonation should solely reside in a rise occurring at the end of the utterance. Frequently, he maintains, questions have also high beginnings (especially on question words), and on the whole they are realized on a higher global pitch level than are statements. To illustrate this, Hermann claims that these two sentence types are typically realised on different pitch levels, i.e. statements roughly between level C and B, questions roughly between level D and A (see Figure 2.3 below). Though it is possible for tones to be realized outside these default ranges, Hermann claims that higher tones in statements will not reach as high as level D, and that lower tones in questions will not become as low as level B. In fact, this seems to be the case in all of the 175 languages studied (Hermann 1942:363).

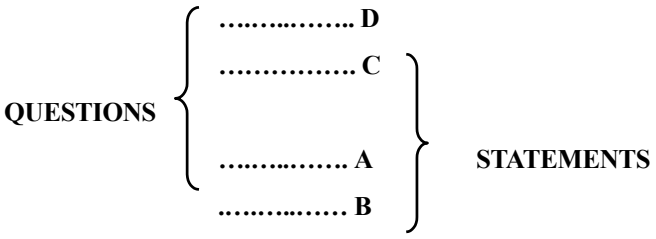


Figure 2.3. Crosslinguistic use of pitch levels in questions vs. statements. While it is possible for the intermediate levels A and C to overlap, the lowest level B is only used with statements, the highest level D only with questions (after Hermann 1942:155)

To Hermann, the fact that this specific 'question pitch' ('Frageton') is found in numerous unrelated languages scattered all over the world, tone languages included, strongly suggests that it stems from a common root. Taking his cue from young children, who produce question pitch on one-word utterances long before they learn to use specific question words or inversion, he proposes that the overall development of human language has gone through the same stages. That is, early man must have felt the need to distinguish one-word informative utterances from one-word interrogative utterances by giving the latter higher or rising pitch. Since, according to Hermann, higher pitch is better perceived it can be regarded as an appeal for extra attention, necessary to get the listener to make some response. As, in these primitive 'questions', high final pitch signalled that the speaker anticipated a reaction, it took on the more general meaning of incompleteness. In later stages of the development of human language, specific question words came into existence which, likewise, were realized on relatively high pitch. The circumstance that such question words were, mostly, utterance-initial would then account for the overall falling intonation of wh-questions

vis-à-vis the rising intonation of yes-no questions. To sum up, in Hermann's account high(er) question pitch springs from a single source and is therefore universal. A remnant of a very early stage in the development of human language, it represents an appeal for attention; secondarily, it expresses incompleteness. If, at times, the 'Frageton' is seemingly absent, closer scrutiny usually causes it to be spotted in some disguise or other (Hermann 1942:363).

Ohala (1983, 1984, 1994), likewise, regards high question pitch as a universal phenomenon. Delving even deeper into the history of man than Hermann, he claims that high question pitch has an ethological basis. He brings in evidence that, in birds and mammals large and small, relatively high pitch is used to signal subordination/submissiveness, as opposed to relatively low pitch signalling dominance/confidence[15]. In face-to-face encounters between two animals of the same kind, these contrasting frequencies are assumed to communicate the individuals' respective sizes (cf. the growling large dog vs. the whining little puppy), which, presumably, will go some way towards avoiding bloodshed. Thus, apart from having the basic meaning 'large', low/falling pitch in a vocalizer is also associated with dominance, confidence and self-sufficiency. High/rising pitch, by contrast, signals smallness of the vocalizer and, by extension, lack of threat, subordination, dependence. According to Ohala, this cross-species, pre-linguistic Frequency Code accounts for a number of systematic sound-meaning correlations in human language. One of these is the cross-language correlation between sentence type and use of pitch, i.e. high/rising pitch in questions vs. low/falling pitch in statements. The Frequency Code would explain why someone asking a question uses high pitch: for this piece of information, the speaker depends on the listener's goodwill and cooperation. If the Frequency Code is an integral part of the human vocal communication system it also explains why, crossculturally, attitudes and emotions such as politeness or fear are, similarly, conveyed by either high/rising or low/falling pitch. As a matter of fact, according to Ohala the use of pitch as motivated by the Frequency Code should be seen as an equivalent of a bodily gesture such as raising one's eyebrows, i.e., it is superimposed on the linguistic message and adds some paralinguistic element to it (1984:5).

Bolinger (1978, 1982, 1989) also places the differential use of pitch in the gestural domain. It is his conviction that there is a universal opposition between 'up' and 'down' which underlies both human bodily gestures and intonation. 'Down' is associated with rest/relaxation, reflecting that pitch naturally reverts to a default low level; secondarily, it has the abstract meaning 'completeness'. 'Up', by contrast, relates to strain/tension. Yet, while the production of high pitch requires an effort of the vocal organs, there is no such thing as a default *upper* level of pitch analogous to the default *lower* level. That is, gradiently higher pitch commonly corresponds to greater emphasis or stronger emotions. Secondarily, high pitch signals 'incompleteness, openness'. On this basic binary opposition of the features UP and DOWN, speakers and hearers are able to build an elaborate pragmatic structure (1982:18). That is, cross-combinations of these two features with different grammatical sentence types yield subtypes carrying

[15] According to Morton (1977), cited in Ohala (1983:7), 28 avian and 28 mammalian species were found to make identical use of the Frequency Code.

different shades of meaning, depending on whether 'upness' or 'downness' has been superimposed. For example, a statement may be rendered 'incomplete' by rising intonation, indicating that the speaker wishes to add something, as may be the case in list intonation. Given its ties to human physiology, the dichotomy UP-DOWN running through intonation is obviously universal. At the same time, there is not a universal intonation for questions, since speakers are at liberty to combine their questions with either of the features UP and DOWN[16]. Nonetheless, when Bolinger dismisses the idea that questions may have a characteristic intonation of their own, he mainly concentrates on *grammatical* questions, i.e. on utterances featuring inversion, a question word or both. However, it has repeatedly been shown that such utterances may have all kinds of illocutionary forces other than questioning (cf. chapter 1, § 1.2.1). For intonation research, it is the *pragmatic* questions that are of interest, i.e. utterances which the speaker *intends* to function as a request for information, irrespective of lexical or syntactic markers. Also, when Bolinger denies questions a specific intonation he mainly refers to the patterning of their pitch accents. At the same time, he also acknowledges that there are universal tendencies for questions to display rising pitch (notably yes-no questions, 1978:501) and to have higher average pitch than statements (1982:6).

The universalist claim that (question) intonation is highly similar in all types of languages because it represents some direct phonetic expression of, e.g., a speaker's frame of mind or physical make-up, also meets with scepticism. According to Ladd (e.g. 1990, 1996), the presence of crosslinguistic similarities in intonation is no reason to deny the existence of a language-specific, underlying phonological level. For instance, when it comes to accent placement in questions different languages may have different rules, causing question contours to have different shapes. Thus, as Ladd points out, there are differences between, for example, Russian, Hungarian, Romanian or Greek question contours on the one hand and, on the other, Dutch, German, Italian or French question contours (1990:815). Such language-specific variation would seem incompatible with a view that (question) intonation directly springs from some universal source.

To sum up, crosslinguistically the defining property of question intonation appears to be high pitch. While, according to some, this high question pitch typically occurs utterance-finally, others claim that it may also manifest itself in earlier parts of the utterance; it may even raise the entire question onto a higher level. At the same time, the postulated high question pitch meets with diverging analyses. Some investigators view it in a **historic** light: they regard it as a relic of the earliest times of human speech (Hermann), or as having originated from ellipsis of the question part with falling intonation (Kretschmer). Others are of the opinion that high question pitch is **iconic**, i.e., it is taken to derive from the physical make-up of man (Bolinger), or from archetypal cross-species behaviour (Ohala). Finally, theorists such as Ladd prefer the view that high pitch in questions is, on the whole, **conventional**. That is, when a question happens to have high pitch this is largely the result of the speaker having chosen certain phonological units which, accidentally, have caused pitch to be high.

---

[16] "It has been emphasized repeatedly in this book that no intonation is an infallible cue to any sentence type: any intonation that can occur with a statement, a command, or an exclamation can also occur with a question" (Bolinger 1989:98).

In an attempt to bring the iconic and conventional approaches together, Lindsey (1985:49) puts forward that it cannot be doubted that intonation has an iconic base, but that it cannot be denied either that it can take on conventional meanings which may be quite language-specific. Thus, in a given utterance it is possible for elements from a relatively iconic source to mix with elements that must be regarded as conventional(ised). However, Lindsey also assumes that it is impossible for such conventional(ised) meaning to clash with basic iconic meaning. Having investigated the tonal expression of interrogativity in English, Hausa, Luganda and Cherokee (specifically of yes-no questions), Lindsey concludes that there is a typological continuum along which languages may be 'intonationally rich' or 'intonationally poor'. The former differ from the latter in that they may employ a wide range of tonal mechanisms for the expression of interrogativity. Thus, in yes-no questions, intonationally rich languages typically make use of (i) high tones rather than low, (ii) a raised rather than a lowered accent peak, (iii) a high final boundary tone rather than a low one (i.e., a final rise), and/or (iv) pitch that is globally raised. That is, like Hermann, Lindsey is of the opinion that high question pitch may manifest itself locally as well as globally.

Taking Lindsey's four parameters for a starting point, it seems useful to look at individual studies of question intonation in a variety of languages, in order to determine what (local or global) means are employed by speakers to express interrogativity. Of these studies, § 2.2.2 provides an overview that, it should be noted, does not pretend to be exhaustive. Next, § 2.2.3 is more specific in that it deals with claims and empirical findings concerning question intonation in Dutch.

2.2.2 High question pitch

2.2.2.1 Local high question pitch: The final rise

> *Parry had his generation's habit of making a statement on the rising inflection of a question - in humble imitation of Americans, or Australians, or, as I heard one linguist explain, too mired in relative judgements, too hesitant and apologetic to say how things were in the world*[17].
>
> From: *Enduring Love,* Ian McEwan (Vintage, 1998)

It cannot be doubted that, across languages, the intonation of questions is frequently characterized by a sharp final rise in pitch. Bolinger (1978:501), pointing out that Ultan (1969) already found a predominance of rising final pitch in 53 languages, goes on to add a further 41 languages. Later, Sadock and Zwicky (1985:181) state that, in their sample of 23 languages, "rising final intonation is one of the most frequently found indicators of interrogative force, [……]". Nonetheless, Geluykens (1988) challenges claims that questions typically end with a rising pitch movement. Having investigated

[17] For this phenomenon, see Guy *et al.* (1986).

216 spontaneous[18] British-English 'polar' questions (i.e. 156 yes-no questions and 60 declarative questions, or 'queclaratives'), he concludes: "There is no justification for regarding rising contours as the unmarked pattern for polar questions" (1988:468), and: "In queclaratives, a falling intonation is by far the most frequent pattern, mostly accompanied by a step-up in pitch in the head of the tone unit. […] In all, the claim that rising intonation (and, more particularly, final Rises) is the 'normal' pattern for polar questions lacks empirical justification" (1988:467). Two things may be noted here.

First, Geluykens's procedure for identifying spontaneous utterances as (declarative) questions would seem somewhat questionable. All utterances followed by what he calls a "typical answer" (i.e. confirmation/*yes,* rejection/*no*, or *perhaps, I don't know*) were classified as questions, i.e. as "genuine requests for information". However, in spontaneous speech, more often than not the frequently occurring monosyllable *yes* does not constitute an answer. It merely represents 'backchannelling', i.e., the listener indicates: "I'm listening, carry on". Thus, e.g. statements or exclamations may equally be followed by *yes*. Obviously, in such cases it would not seem appropriate to automatically label the utterance as a question[19]. Indeed, some of Geluykens's examples of queclaratives suggest that they actually have the illocutionary force of an exclamation or a conclusion, rather than of a 'genuine request for information'. Thus, if the author observed rising pitch in only 52% of the yes/no questions and 43% of the queclaratives, this selection method should be borne in mind. Also, as the author himself points out, it is quite possible that the questions without rising tones might have carried attitudinal overtones rendering them less prototypical.

Second, if it is claimed by Geluykens that a good many questions end in a falling nuclear tone rather than a rising, the author also shows that, in the selected questions, there was a marked tendency for pitch preceding the nuclear syllable to be raised. This was especially so in the queclaratives, whose question status depends entirely on intonation. The effect of these local upsteps in pitch was to raise the f0 of the starting point of the nuclear syllable and cause the accentual fall to start relatively high in the speaker's pitch range. Moreover, the global pitch levels in Geluykens's questions also tended to be high in the speakers' ranges. In effect, his findings strikingly corroborate Lindsey's conclusion that, in intonationally rich languages, there are more high-pitched cues to interrogativity than just a rise at the end (see also Geluykens 1989:574). Such cues will be dealt with in the next section.

2.2.2.2 Local high question pitch: Other manifestations

In a number of languages, questions are reported to have higher onsets than the corresponding statements. Thus, in Finnish, both questions and statements lack final rises and display overall declining pitch, but in the questions, initial pitch is higher (Iivonen 1998:318). In Western Arabic (spoken in Morocco), statements and yes-no

[18] It may be pointed out that there are indications that question intonation in spontaneous speech behaves differently from question intonation in read speech (e.g. in Bari Italian, cf. Refice, Savino & Grice 1997).

[19] The same procedure of selecting pairs of 'questions' and 'answers' (termed 'adjacency pairs') was followed by Brown, Currie & Kenworthy (1980:170, 180) who, (accordingly?), did not find *any* declarative question with rising intonation. Selting (1992), also, used this method. For a critique, see Hetland & Molnár (1995).

questions share a rising-falling pattern, but again, onsets of questions are higher (Benkirane 1998:353). Similarly, interrogativity may cause accented and unaccented peaks to be higher than in the corresponding statements. Thus, for instance, in Swedish, Dutch and Western Arabic questions the peaks of the nuclear accents are raised (Gårding 1983:23, Hadding-Koch & Studdert-Kennedy 1964, Van Es 1932, Benkirane 1998:354). In Bengali, the statement and question version of the segmentally minimal pair *That boy is TALL* vs. *Is that boy tall?* both end in a rising-falling movement. However, in the question version the f0 maximum of the rise-fall is considerably higher than in the statement version (even when taking into account that, in Bengali questions, overall pitch range tends to be wider, Hayes & Lahiri 1991:65). More in general, pitch range proves particularly sensitive to interrogativity. In Kiswahili, the intonational contrast between statements and questions corresponds to a low rise vs. a high rise, respectively (Maw & Kelly 1975, as cited in Hetland & Molnár 1995:23). In Danish, which is claimed to have a single, recurring basic stress group rather than an inventory of different pitch accents (cf. Grønnum 1992:49), the (unaccented) high tones in these groups rise higher in questions than in statements (Grønnum 1995:128). Hausa, a tone language, makes additional use of pitch to encode interrogativity. One of its devices in yes-no questions is raising the last syllable in the utterance bearing a lexically H tone to a superhigh level (Lindsey 1985:105; Inkelas & Leben 1990:18). Similarly, in Chengtu Chinese, different tonemes (i.e. contrastive tones) on the final syllables of questions are all somehow raised in pitch. For instance, in questions the low-falling toneme becomes low-level, and the low-falling-rising toneme loses its fall and becomes low-rising (Chang 1958, as cited in Ladd 1996:150). In the Dutch (Limburgian) dialect of Roermond, the only difference between a statement and a wh-question is in the starting pitch of the final rise: in the former, this is low, in the latter high (Gussenhoven 1999b:290).

Apart from the local final rise or the local raising of (final) peaks, lexical tones or tonemes, it is also possible for interrogativity to be encoded by using a specific *type* of nuclear pitch accent. Often, this is rising, as opposed to a non-rising nuclear accent in statements. This is the case in, e.g., Neapolitan Italian, which has a rising pitch accent for a question, a falling for a statement; there is no final rising boundary tone (D'Imperio 1997). Grice (1995), Grice *et al.* (1995), and Grice & Savino (1997) found that, in some varieties of German, Italian and Bulgarian, interrogativity is signalled, not by a final high boundary tone but by a (prefinal) rise which should be seen as part of the accent. Thus, although questions have a *low* final boundary tone (L%), they also have a *rising* nuclear pitch accent, e.g. L*+H or L+H* (where the starred tone associates with the accented syllable)[20]. Similarly, in Hungarian, Greek and Romanian, questions have a *high-falling* final boundary tone (HL%)[21] preceded by a *low-rising* pitch accent (Gósy & Terken 1994; Ladd 1996:115, 212, Mennen 1999:76, 83, Gussenhoven 2000b). The same is reported of the southern Dutch dialect of Roermond, where statements have a

[20] For instance, while Bari Italian statements have the tonal sequence H+L* L-L%, in (spontaneous) yes-no questions it is L+H* L-L%.

[21] Note, however, that in Greek and Romanian the exact position of the H of HL% depends on the location of the nuclear accent. When this does not occur on the last word but earlier, the H of HL% will associate with the stressed syllable of the last word (Ladd 1996:213, Mennen 1999:76).

high/falling accent (H*) followed by a low boundary tone (L), whereas questions feature a low/rising accent (L*) together with a high-falling boundary tone (HL; Gussenhoven 2000a:134).

These examples show that, in a number of languages, interrogativity may be phonologically encoded by pitch accent type and boundary tone in tandem. If one of them is phonologically low or falling, the other is high or rising, and vice versa. This division of labour, which appears to be more or less fixed, would also seem to hold in the phonetics: as perception experiments have shown, several phonetic cues to questionhood may interact (Hadding-Koch & Studdert-Kennedy 1964, and Studdert-Kennedy & Hadding 1973). Thus, when in one of these potential cues (e.g. the accent peak) pitch is comparatively low, it has to be truly high in another for the utterance to be perceived as a question. The next section further elaborates on such instances of compensatory pitch relations.

2.2.2.3 Local high question pitch: Interactions

With an explicit reference to Hermann (1942), Hadding-Koch & Studdert-Kennedy (1964) systematically varied the f0 values of accent peak height (2x), postaccentual low ('turning point', 3x) and offset (7x) of the Swedish utterance *För Jane.* Subjects had to indicate whether the utterance was a statement or a question, and it was found that question responses increased as a function of an increase in f0 at *any* of the three variable points. For instance, a final fall of a certain size got 96% statement responses when the turning point was low, as opposed to 89% question responses when the turning point was high[22]. The authors concluded that it is shape and level of the *entire* configuration of accent and subsequent boundary tone which determine whether or not an utterance is perceived as a question.

In a later experiment, Studdert-Kennedy & Hadding (1973) extended their investigations, in an attempt to determine how the various pitch cues to questionhood are weighed by listeners. Again, a short utterance (*November*) was varied as regards the f0 of peak (3x), turning point (i.e. the postaccentual low; 4x) and offset (6x), and subjects had to judge whether an utterance was a question or a statement (linguistic judgements). However, there were two further tasks. The same utterances were also presented as frequency-modulated trains of pulses and as sine waves, and in both cases the subjects had to decide whether they heard a final rise or fall. It was found that, first, the terminal movement of f0 was the single most powerful determinant of the 'question' vs. 'statement' judgements. That is, none of the contours which in 90% had been classified as questions lacked a terminal rise, and few of their statement counterparts lacked a final fall. Second, for a correct classification of the final movement as either a fall or a rise, the relation with pitch level and pattern of the preceding portion of the contour seemed to play an important and complex role. On the one hand, rises and falls following an onset of 130 hz and a turning point of 145 hz (i.e., a more or less level

[22] O'Shaughnessy (1979:141), comparing this perceptual effect of synthetically raised turning points with his own natural production data of American English questions, speculates that "a high f0 value at that crucial point serves the same f0 cue as generally-high f0 in natural Qs [i.e. questions]".

stretch) were quite accurately perceived. However, when the turning point was higher (160 hz) judgements varied. When the turning point was followed by a steep fall (from 160 to 100 hz), listeners needed a final rise of 50 hz to classify the final movement as a rise. But when, following the same turning point of 160 hz, pitch kept on rising to 180 hz, even final *falls* of 50 hz. were classified as rises. Apparently, this steadily rising pitch was strongly suggestive of questionhood; whether the final movement following it was actually rising or falling seemed immaterial. The height of the turning point also affected the judgements in the sine wave and pulse train conditions, which suggested that the effect was primarily auditory. According to Studdert-Kennedy & Hadding, the f0 of the turning point probably influences linguistic judgements indirectly rather than directly, in that variation alters listeners' perceptions of the final pitch movement. By contrast, peak height specifically affected the linguistic judgements, suggesting that it has a direct linguistic function. The authors concluded that the perceived direction of the final movement is being weighed together with peak height. When listeners hear a final rise, they interpret the utterance as a question unless the rise is only slight and the peak unusually low. When listeners hear a final fall, the utterance is interpreted as a statement, unless the fall is slight and the peak is very high (Studdert-Kennedy & Hadding 1973:310). Thus, paradoxically, though the direction of the final movement would seem decisive it is not always correctly perceived: judgements were clearly influenced by level and pattern of preceding pitch. In particular, an utterance with gradually rising pitch proved prone to being perceived as a question, even in the absence of a final rise. In short, listeners appeared to process the contour of *November* as a unit, constructing questionhood 'on-line', i.e., on the basis of pitch relations between the different components of the contour.

Such a view of questionhood being wholly or partly cued by (semi-local) pitch relations seems to tie in with production data. In spontaneous yes-no questions in Palermo Italian, interrogativity is signalled by the *transition* between the prenuclear part of the contour (the 'head') and the following nuclear tone (Grice 1995:33ff). In this variety of Italian, both the statements and questions may end in a nuclear fall from mid or high pitch. A question interpretation depends on the pitch level immediately *before* the starting point of this fall. When this is relatively low, so that pitch has to *rise* in order to reach the level from where it finally starts to fall, the utterance has question status. If, however, the prenuclear part of the contour ends in the same or higher pitch than the starting point of the nuclear fall, the utterance is a statement. The effect of this is that, in statements, the final part of the contour thus has an overall falling shape, whereas in questions it is rising-falling. Similarly, Geluykens found (cf. § 2.2.2.1) that both the yes-no and declarative spontaneous questions in his corpus of British English frequently combined a final fall with a slowly rising 'head' (i.e. the portion of the contour preceding the nucleus).

All in all, languages appear to display a variety of phonological and phonetic devices for encoding interrogativity, and all of these somehow involve high(er) or rising pitch. Locally, a final rise and/or a raised (accent) peak can serve the purpose, as can a specific pitch accent. Also, interrogativity may manifest itself in specific pitch relations among different components of the contour. As a matter of fact, this array of possibilities would seem to sit somewhat uneasily with a compositional approach to

intonational meaning (cf. § 2.1.3). This theory claims that pitch accents and boundary tones each have scope over different domains and convey specific information on the way in which the listener is supposed to interpret the utterance[23]. This seems hard to reconcile with the above observations that both categories of tone can be employed to express the function of interrogativity. This, in turn, might be taken to mean that the division of tones into pitch accents and boundary tones is not always as clearcut as one would wish (cf. Ladd 1996:213; see also Grice 1995).

2.2.2.4 Further functions of the final rise

Returning once more to the final rise it may be mentioned that, in some accounts, this rise is not regarded as automatically expressing interrogativity or, for that matter, continuation. Selting (1992:340), who views intonation as an autonomous signalling device for the organisation of conversational interaction, claims that speakers do **not** use the final rise (in German wh-questions) to convey the interrogative status of the utterance. Instead, this rise has the function of distinguishing between different types of conversational questions. Thus, a rising final pitch movement is taken to indicate that the question is 'unrestricted', i.e., it introduces a topical focus that is new in relation to the preceding conversational turn. By contrast, with falling final pitch the question is taken to be 'restricted', i.e., the speaker asks for information that is related to a previously established topical focus.

Similarly, in Edinburgh Scottish English, the final pitch of spontaneous questions is claimed to distinguish between question types (Brown, Currie & Kenworthy 1980:190). Edinburgh questions are observed to end with a rise, with a fall-to-mid-pitch or with a fall-to-low-pitch. According to Brown *et al.*, the final f0 value reflects whether the question is conducive or non-conducive, that is, whether or not the speaker has the expectation of one answer being more likely than the other. In fact, this would seem to amount to the difference between questions for confirmation and questions for information, respectively. Now the Edinburgh questions with high or mid final f0 values are claimed to be non-conducive, whereas those with low endings are felt to be conducive. Similarly, it is claimed that French yes-no questions are marked by a final rise only when they seek information; falling pitch indicates that the speaker wants confirmation (Di Cristo 1998:203)[24]. In chapter 1 (§ 1.3.1), it was pointed out that confirmation questions require fewer bits of information and would therefore seem less prototypical than information questions. Apparently, it is possible for this distinction to

[23] Pierrehumbert & Hirschberg (1990:308): "[….] pitch accents convey information about the status of discourse referents, modifiers, predicates, and relationships specified by accented lexical items', and "[…] boundary tones convey information about the directionality of interpretation for the current intonational phrase - whether it is "forward-looking" or not''.

[24] The distinction between questions for information or confirmation may also be signalled by accent type. According to Gussenhoven (1984a:203/1988), Dutch information questions are cued by an accentual rise (L*H), confirmation questions by an accentual fall-rise (H*LH). Similarly, in Bari Italian questions for confirmation ('checks') have rising pitch accents (L*+H) as long as they are tentative (and hence resemble 'queries' which seek new information), but falling accents (H*+L) when the speaker feels confident that old information will be confirmed (Grice & Savino 1997:33).

be intonationally encoded, for instance by the presence/absence of a final rise. We will return to the issue of information questions vs. confirmation questions in chapter 5.

Finally, high final pitch has, of course, been also widely claimed to signal 'continuation' which, in turn, would seem to rest on the more abstract meaning 'openness' (e.g. Cruttenden 1981, 1994; Hirst & Di Cristo 1998:27). For instance, when an utterance is comprised of two or more grammatical units, non-final units are bound to end in high/rising pitch, indicating that the speaker has not finished yet. More about this particular function will be said in § 2.2.3.

2.2.2.5 High question pitch: Global

2.2.2.6 Raised register level

Although a final rise seems a major diagnostic of questionhood, yet it occurs fairly late in the utterance. Considering that a question makes an appeal to the listener to come up with an answer, it would be advantageous if high question pitch were present earlier, enabling a listener to process the utterance as a question right from the start. It is not surprising, therefore, that in many languages questions are found to be realised on a higher register level[25]. That is, all tones, whether H(igh) or L(ow), accented or unaccented, are raised relative to their f0 values in corresponding statements. The effect seems especially conspicuous in declarative questions, whose question status crucially depends on intonation. Examples of raised register level would be Swedish (Hadding-Koch & Studdert-Kennedy 1964:175, Gårding 1983:21), Edinburgh Scottish English (Brown, Currie & Kenworthy 1980:59); Cherokee (Lindsey 1985:144), British English (Geluykens 1988:477), Hausa (Inkelas & Leben 1990:18), American Spanish (Sosa 1991:119), Brazilian Portuguese (de Moraes 1998:184), Finnish (Iivonen 1998:319), Western Arabic (Benkirane 1998:354), Mandarin Chinese (Shen 1990:25), Chichewa (Myers 1999), Chickasaw (Gordon 1999), as well as American English, Hungarian, Vietnamese and Thai (Hirst & Di Cristo 1998:25)[26].

2.2.2.7 Suspension of downtrends

[25] In the present study, we adopt Clements (1990:59) definitions of the (related) notions *Range* and *Register*: "*Range* refers to the largest interval (or frequency band) within which tones are normally produced in the speech of a given speaker. *Register* is a smaller interval or frequency band internal to the speaker's range, which determines the highest and lowest frequency within which tones can be realized at any given point in an utterance." That is, a speaker's register will usually span only part of his overall range; we refer to this dimension as the 'register span' (cf. Ladd 1996:261). In addition, within the overall range the position of the register (i.e., 'register level') may be shifted upward or downward.

[26] The upward register shift in questions may have been behind the Spanish practice of employing an inverted question mark *before* a question. This special punctuation mark was introduced by the Real Academia Española in 1754 because the single, sentence-final question mark was felt to be inadequate, "lacking indication, […], as to where the interrogative tone begins and continues until it is completed […]" (Parkes 1992:56). The initial question mark effectively alerts the reader to the utterance's question status right from the start and enables him to adjust his intonation.

Another global manifestation of high(er) question pitch is the suspension of downtrends, i.e. of declination, of final lowering and of downdrift. To start with the latter, the phenomenon downdrift is characteristic of tone languages and causes a lexically high tone (H) to be lowered whenever it is preceded by a low tone (L; cf. Lindsey 1985:103; Inkelas & Leben 1990:22). Examples of languages where downdrift is suppressed in questions would be Hausa (Inkelas & Leben 1990), Zulu (Laughren 1984 as cited in Clements 1990:69), and Chichewa (Myers 1999).

The phenomenon 'final lowering' may perhaps be called semiglobal rather than global, as it only involves the lowering of the latter portion of the utterance. It has been demonstrated that, in Japanese questions, this lowering is suspended. Thus, while the latter half of a Japanese statement shows a clear downward trend, this cannot be observed in a question (Pierrehumbert & Beckman 1988:73).

'Declination', finally, refers to the gradual, time-dependent lowering of f0 in the course of an utterance or a text (e.g. Cohen, Collier & 't Hart 1982; Liberman & Pierrehumbert 1984; Thorsen 1985, 1986; Van Den Berg *et al.* 1992; Sluijter & Terken 1993). It is still a matter of dispute whether or not declination is under a speaker's active control. Nonetheless, there are claims to the effect that the steady lowering of f0, while being characteristic of statements, fails to take place in questions, e.g. in French and Russian (Vaissière 1983:57), in German (Oppenrieder 1988:199), and in Danish (Thorsen 1980, 1995). This would seem to suggest that global declination (or its absence) has (acquired) some element of meaning, regardless of whether it is actively controlled. The issue of the 'meaning' or communicative function of declination is further elaborated in the next section.

2.2.2.8 Declination vs. downstep

For some proponents of 'superposition' or 'overlay' models of intonation (cf. § 2.1.1.2), the slope of declination has a linguistic function of its own. This may be illustrated on the basis of Thorsen/Grønnum's analysis of Standard Danish (e.g. 1980, 1992, 1995). This language has a single, recurrent stress group which consists of a brief initial fall on the stressed syllable, optionally followed by a steep unaccented rise and a steep unaccented fall. In statements, a line connecting the stressed syllables will reveal a globally downward slope which, in typical usage, is fairly steep. By contrast, in declarative questions this global trend is level. In between these two extremes, there is a continuum of intermediate slopes[27]. Apparently, there is a trade-off between lexical or syntactic information and the slope of declination: when questions are marked by a question word and/or inversion, their slopes are typically steeper than when they lack such markers. Given the observed variation in declination slopes, Grønnum (1995) argues that, in Danish, intonation cues to speech act are global, not local, and that global

[27] It is worth noting that Figure 2.4 is a schematic representation: the regular distances between the slopes of the various utterance types should be seen as idealised. Thorsen's important finding was that the slopes showed a clearly observable order ranging from terminal statements to questions otherwise unmarked for interrogativity, and that this order was not affected by variation among speakers.

slope, being a linguistic element, must be included into a model. Figure 2.4 illustrates her findings.

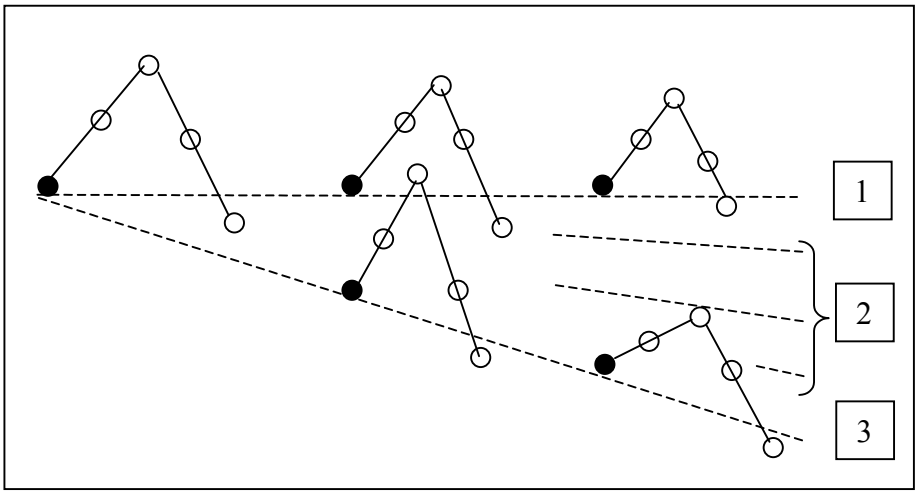


Figure 2.4. Different declination slopes in different utterance types in Standard Copenhagen Danish. Filled dots represent stressed syllables, unfilled dots unstressed syllables. The uninterrupted lines between the dots represent the recurring f0 pattern of Danish stress groups. The interrupted lines indicate the declination slopes of (1) syntactically unmarked questions; (2) questions with inversion and/or interrogative particle, as well as continuatives; (3) statements (after Thorsen 1980).

That declination is perceptually relevant was demonstrated by Gooskens & Van Heuven (1995), who investigated its importance for the recognition of utterance type by Danish and Dutch listeners. It was found that, for the Danish listeners, declination slope constituted an important cue. Ideally, Danish statements should have a fairly steep downward slope, whereas in declarative questions, declination should be absent. On their part, the Dutch listeners made it clear that Dutch questions must have a final rise, a feature which is absent from Danish question intonation; with respect to declination, the results were slightly less clear-cut. Yet, on the whole the Dutch listeners also appeared to favour questions *without* declination, as opposed to statements and non-final clauses *with* declination.

At this point, it must be mentioned that proponents of the autosegmental theory have analysed the *global* declination of f0 as the product of a strictly *local* lowering of accent peaks. That is, each successive peak is claimed to be lowered, stepwise, by some fixed amount relative to the f0 of the preceding peak; this process is termed 'downstep'[28] (e.g. Bruce 1977, 1982, Pierrehumbert 1980, Liberman & Pierrehumbert

[28] Accounts of how downstep is actually triggered vary, though. In Pierrehumbert (1980), downstep is the result of a phonetic realisation rule: after a H*+L pitch accent, the H of the following pitch accent is automatically downstepped. In Beckman & Pierrehumbert (1986), all bitonal pitch accents are claimed to trigger downstep. By contrast, Ladd argues that downstep is an independent speaker choice with an

1984, Van Den Berg *et al.* 1992, Prieto *et al.* 1996). Crucially, this alternative analysis leaves intact a key idea of autosegmental theory, viz. that an intonation contour is a sequence of strictly local phonological events; there is no question of speakers independently choosing a global slope. Yet, advocates of a downstep analysis do recognise the existence of some time-dependent declination component (Beckman & Pierrehumbert 1986:302, Pierrehumbert & Beckman 1988:89, Gussenhoven 1988:10, Gussenhoven & Rietveld 1988). However, while Pierrehumbert & Beckman incorporate the possibility of a coexistence of downstep and declination into their model of Japanese, Ladd (1996:288 n.13) points out that the issue is by no means settled, given the empirical difficulties of testing such a model. As a side-effect of the numerous discussions on 'declination' vs. 'downstep', the terms have come to be associated with the competing theorical frameworks, i.e. the superposition models and the autosegmental models, respectively (for overviews, see Ladd 1984, Nolan 1995, Grice 1995).

2.2.3 Question intonation in Dutch

Of Dutch (question) intonation, it has been claimed: "There is no specific intonation pattern characteristic of questions" ('t Hart 1998:103, cf. § 2.2 above; see also Van Den Berg 1972:106). This, however, is not the only view of Dutch question intonation that has been put forward. Earlier, in 1914, Van Alphen indicated that Dutch questions crucially end with a final rise on the last syllable. This high final pitch reflects the incompleteness of the utterance and it occurs in wh-questions, yes-no questions and declarative questions; it is absent from rhetorical questions, examination questions and alternative questions. Basing himself on musical notations, Van Alphen further concluded that Dutch questions are realised on a higher overall register level than are statements. In his account, asking a question always involves 'tension', which is being caused by the speaker's doubt or by some gap in his information. This tension then manifests itself as a general raising of the pitch throughout the utterance[29].

Zwaardemaker and Eijkman (1928:283), likewise, claim that Dutch questions distinguish themselves by a relatively high register level. Also, while yes-no questions generally have a rising accent followed by a fall and a further rise on the final syllable, wh-questions allow more variety.

Guittart (1925:41, see § 2.1.2 above) noted a mainly falling intonation in wh-questions, as opposed to a predominance of rises in yes-no questions and declarative questions. In general, questions tend to have high initial pitch, brought about by an accented wh-word or by a non-accentual high initial tone.

In his early experimental approach of intonation, Van Es (1932, see § 2.1.2 above) recorded the intonation contour of the neutral question *Komt Piet vandaag ook?*

---

abstract meaning of its own: pitch accents may be either downstepped or non-downstepped. In his account, downstep results from a higher-level phonological pitch range relation (Strong-Weak) between two subsequent pitch accents (Ladd 1996:278). In Gussenhoven's system, downstep is proposed to be morphologically triggered, that is, a speaker optionally attaches the morpheme [downstep] to the utterance (Van Den Berg *et al.* 1992:341).

[29] Note the similarity with Bolinger's (1978) later claim that, in the binary opposition between UP and DOWN, the former involves (bodily) strain and tension (cf. § 2.2.1).

('Does Pete come too, today?') and concluded that the contour had a 'synthetic' shape. That is, although f0 reached its highest point in the vowel of the accented subject *Piet,* the relatively high pitch of the beginning of the utterance had roughly the same level as that of the end. Van Es also compared the question with its conditional/continuative version *Komt Piet vandaag ook, dan blijf ik thuis* ('Should Pete come too, today, then I'll stay at home'). Interestingly, though both versions ended more or less on the same high f0 value, the question featured a clear-cut final rise in the last syllable, whereas the continuative version roughly maintained a high pitch level from the accent onwards. Besides, Van Es observed that, in the question version, the f0 of the peak was markedly higher than in the continuative clause. This, he argues, causes it to contrast with the rest of the utterance. Subsequent acoustic analyses of more questions of different types led Van Es to conclude that, on the whole, Dutch question intonation is characterised by a high beginning and a raised nuclear accent peak, followed by a distinct valley (the 'interrogative valley') and a sharp rise in the final syllable. Phonetically, this rise differs from the accentual rise in that it does not make the syllable/word prominent. Van Es emphasises that defining question intonation as merely consisting of a final rise fails to take into account the roles played by other local and global characteristics, as well as by their mutual relations (Overdiep & Van Es 1949:88, 101). This, obviously, anticipates the later findings of Hadding-Koch & Studdert-Kennedy discussed in § 2.2.2.3 above. Daan (1938:476), too, measured pitch in a number of statements and questions. Like Van Es, she also came across the characteristic 'interrogative valley' or 'closed form' of questions identified by him.

As Van Es's comparison between an interrogative and a continuative contour shows, high final f0 may appear in at least two different shapes. First, taking off from the baseline pitch may keep on rising all the way up (the all-rise variant). Second, when pitch is already high after an accentual rise, it may not revert to the baseline but continue on a high level course (the high-level variant). In point of fact, as shown by Caspers (1998) this formal difference also plays a role in the process of taking turns in conversation. In a perception experiment Caspers found that, at the level of discourse, the high-level variant generally functions as a cue to turn-keeping ('please do not interrupt'). The all-rise variant may signal the same, but it may also communicate that the utterance is a question, in which case the speaker yields his turn to the hearer. Although, at first blush, this might seem contradictory it is not. Obviously, what the all-rise and high-level variants have in common is the high-pitched end-point, whose underlying abstract meaning is 'incompleteness' or 'more is to come' (see also Keijsper 1984:123). That is, the utterance may sound incomplete either because the speaker will carry on speaking, or, in the case of a question, because the hearer is expected to come up with an answering statement. Still, in the question contexts Caspers' subjects clearly preferred the all-rise contour, whereas the high-level contour was preferred in continuation contexts. It is not immediately obvious whether this preference for the all-rise in the questions was mainly due to the rising movement (H%) *per se*, or to the fact that the (naturally produced) stimuli with the all-rise ended in a substantially higher f0 than the high-level stimuli (a difference of roughly 60 hz., in the male as well in the female speaker).

Finally, as regards the occurrence of declination in questions 't Hart (1998:100) maintains that, in Dutch dialogues, there is no evidence that declination slope depends on utterance category (statements vs. questions), as proposed for Danish. However, this claim would not seem fully in line with the finding of Gooskens & Van Heuven (1995) that Dutch listeners prefer Dutch questions *without* declination.

Taking stock, we can say that the following local and global acoustic properties have so far been attributed to Dutch questions:

(i) a high beginning;
(ii) a relatively high nuclear accent peak;
(iii) a final rise;
(iv) a globally raised register level;
(v) absence of declination.

By and large, these five properties are claimed to be characteristic of yes-no questions (YQ) and declarative questions (DQ). As for wh-questions (WQ), according to some reports these would seem to have a different, overall falling intonation pattern.

2.2.4 Question intonation: Summary

Summing up the previous three sections, we can say that there appears to be plenty of crosslinguistic evidence for the claims of Hermann (1942) and Lindsey (1985) that there is a certain latitude in how interrogativity is intonationally encoded. Crucially, this freedom would seem to be restricted by the basic requirement that, in questions, pitch should be somehow higher than in statements. Thus, what questions across languages have in common is the strong tendency to feature raised pitch levels in onsets, peaks, starting points of rises, end points of falls, and in lexical tones.

Although the feature 'high(er) question pitch' may impress one as fairly vague (cf. Ladd 1996:115), its striking predominance in related and unrelated languages, in tone languages as well as in non-tone languages, suggests a firmly iconic basis. This is not to say that languages cannot deviate from iconicity by making use of conventional(ised) forms. Whereas, in the view of Lindsey (1985:147), it is not possible for a process of conventionalisation to violate basic iconic meaning, Gussenhoven (2002a, to appear) demonstrates that such violations do in fact occur. Importantly, however, in such cases speakers may still exploit the process of phonetic implementation to give expression to universal/iconic meanings. In chapter 5, the issue of the status of universals in intonation theory will be discussed at greater length.

As regards possible characteristics of question intonation in Dutch, opinions appear somewhat divided. While some analysts have claimed that there is no such thing as a specific question intonation, others have pointed out that, in questions, the course of f0 strikingly differs from that in statements. Thus, questions are claimed to display a higher beginning, a raised nuclear accent peak, a final rise, a higher overall register level and/or no declination. The main aim of the present study is to systematically verify these claims by means of a comprehensive production experiment.

2.3 Sex of Speaker

In the literature, it has been repeatedly asserted that male and female speech differ in a variety of respects. However, it has also been pointed out that, of old, opinions on the subject have been powerfully coloured by stereotyped ideas. Accordingly, it is sometimes difficult to tell fact from fiction (cf. Henton 1989). Nevertheless, with respect to a number of sex-related speech differences the results of empirical studies seem to converge. Thus, apart from being higher pitched for anatomical as well as cultural reasons (Van Bezooijen 1995), female speech has also been shown to be more involved, more listener-directed and more standard than male speech. As regards the intonation patterns of females, these have been claimed to be more varied and more expressive (for overviews, see Berryman-Fink & Wilcox 1983, Meeuwisssen 1988, Henton 1989). It would not seem implausible for the latter two properties to be encoded by a wider register span, provided that f0 is measured on a non-linear scale permitting a straightforward comparison of male and female speech[30]. 'More varied' may be taken to mean a greater variety of choices from the inventory of possible pitch configurations, but also, in acoustic terms, greater variation around the average pitch. 'More expressive' may be expected to show up as larger pitch intervals and hence as, again, greater variation around the average pitch.

A subsidiary aim of the present study is to investigate whether the speech of female speakers displays a systematically wider register span than that of the males. Specifically, we expect possible intonational differences between the sexes to be greater in *questions* than in statements, considering the additional claim that men and women have fairly specific communicative patterns. Thus, while the communicative behaviour of men is generally directed at 'agency', that of women mainly aims at 'communion'. As questions are a highly suitable means of establishing interaction, one would expect women to feel more inclined towards asking questions. In addition, women appear to be less afraid than men to show the dependence implied by asking questions; for them, this may form part of a strategy to achieve aims or get things done (e.g. Meeuwissen 1988, Basow 1992). If it is generally the case that women are more positively inclined towards asking questions, it seems not unreasonable to assume that they will make full use of intonational devices for expressing interrogativity, more so than men. On the assumption that these devices include locally widened register span (resulting from larger nuclear accents), as well as a final rise, we formulated the following hypothesis with respect to the independent variable SEX OF SPEAKER:

- Assuming that, in female speech, the implementation of intonational features of questions will be more pronounced than in male speech, we expect the women to produce larger nuclear accents and larger final rises than the men.

2.4 Speaking Style

[30] In chapter 3, it is argued that the most appropriate scale here is de semi-logarithmic ERB scale (Equivalent Rectangular Bandwidth; Hermes & Van Gestel 1991).

The study of some aspect of intonation requires a set of carefully controlled data. Usually, this means that use is being made of read speech, despite comments that this speaking style is not 'natural' enough (e.g., Brown, Currie & Kenworthy 1980:17; Oppenrieder 1988:173). Ultimately, any conclusions based on read speech should be put to the test in unpremeditated speech to check their validity and robustness. In point of fact, for the purpose of the present study we set up a data base of 234 spontaneous questions, taken from doctor-patient interactions (cf. Van Heuven, Haan & Pacilly 1998[31]). However, due to poor sound quality, speakers' overlap and an imperfect numerical balance of the four utterance categories ST, WQ, YQ and DQ, the spontaneous corpus was no match for the corpus of read speech. Hence, it was hard to make straightforward comparisons between the spontaneous material and the material obtained in the laboratory setting. However, considering that the latter yielded more than enough information for a first, exploratory description of Dutch question intonation, there seems no harm in deferring the study of spontaneous questions till future research.

2.5 Hypotheses

Throughout the present study, the acoustic features of questions are compared and contrasted with their 'default' counterparts in statements (ST). On the basis of the various claims provided in the literature, we expect the questions to feature:

(i) a higher-pitched beginning;
(ii) a final rise;
(iii) a raised nuclear accent peak;
(iv) a globally raised register;
(v) less declination.

These five hypotheses will be tested in the three question types selected for investigation in chapter 1:

- wh-questions (WQ), lexico-syntactically marked by question word and inversion;
- yes-no questions (YQ), syntactically marked by inversion;
- declarative questions (DQ), lexico-syntactically unmarked,

Over and above the hypotheses (i-v), we formulate a Functional Hypothesis (FH). Considering that (i-v) predict that the questions generally feature *higher* pitch than the corresponding statements, the Functional Hypothesis expects the three question types to display this high(er) pitch in different *degrees*. Specifically, the FH predicts that high question pitch will be maximally present in questions that are not otherwise marked for interrogativity (DQ), somewhat less in questions with inversion (YQ), and least in questions with both a question word and inversion (WQ). In other words, the FH expects a functional trade-off between syntactic and/or lexical markers of interrogativity

[31] We are grateful to the Netherlands Institute of Primary Health Care (NIVEL, Utrecht) for making (depersonalised) recordings of these interactions available to us.

and high(er) question pitch. This would then result in the order DQ > YQ > WQ, where '>' reflects 'higher-pitched than'. Finally, we expect the female speakers to display a wider (local) register span than the male speakers in, notably, questions.

2.6 General Summary

Question intonation, as it is understood here, is the intonation typical of utterances meant to function as questions, no matter whether or not they are marked for interrogativity by lexical or syntactic features. Across languages, the intonation of questions is commonly found to differ from the intonation of statements. The first part of the present chapter discussed intonation and some of the problems intonational theory has (had) to cope with. Different methods for investigating intonation and identifying descriptive units were briefly reviewed in § 2.1.1. Next, the main features of the two current models of Dutch intonation were presented (§ 2.1.2). Section 2.1.3 made it clear that it is by no means easy to link intonation to linguistic meaning or function. It was observed that intonational contrasts are less clear-cut than segmental contrasts, but that it is nevertheless possible to link certain formal categories to certain broad meanings, notably in the area of how a listener is intended to process the given information (discoursal meaning).

The second part of the chapter was more specifically concerned with the intonation of *questions*. Section 2.2 briefly explored whether there is such a thing as 'question intonation' and whether, by itself, this intonation is capable of distinguishing between statements and questions. Section 2.2.1 reviewed various proposals respecting the origin of the high/rising pitch typically found in questions. It emerged that, from Hermann (1942) onwards, the narrow yet tenacious view that question intonation is identical with a final rise has been gradually replaced with a more subtle outlook. What seems to distinguish question intonation, across the world, is high(er) question pitch occurring **somewhere** in the utterance. This question pitch may occur locally as well as globally and it may take varying shapes. While the subsections of § 2.2.2 focussed on local and global occurrences of high(er) question pitch across languages, § 2.2.3 more explicitly dealt with claims regarding question intonation in Dutch. By and large, a Dutch question contour has been said to have a high-pitched beginning and a rise at the end. In addition, the peak of the (nuclear) accent may be raised, as may the register level of the entire utterance. Also, a perception experiment suggested that questions without a declining f0 ('declination') are favoured over versions featuring a gradual downtrend. On the basis of these claims, § 2.5 presented five hypotheses about intonational properties of Dutch questions; all of these involved high(er) pitch.

In addition, a so-called Functional Hypothesis predicted that the three question types selected for investigation would differ as regards the presence of high(er) question pitch, following the order WQ<YQ<DQ. Finally, on the basis of claims that female speech is more expressive and more varied than male speech, and that the communicative behaviour of women is directed towards 'communion' rather than towards 'agency', female speech was expected to display a wider register span, notably in questions.

Chapter 3

Production (1): Question intonation in read speech

3.1 A corpus of read speech: Introduction

Investigating question intonation in a particular language means focussing on a subset of the possible utterances of that language, i.e. the utterances with the discoursal function 'interrogativity'. Identification of intonational characteristics of this subset is best achieved by means of a two-step process. First, a controlled corpus is compiled with a view to comparing questions with their non-interrogative counterparts, i.e. with statement versions of the same sentence. For this goal, read speech is indispensable because it allows sentences to be systematically varied, guaranteeing maximally clear-cut results. When the production corpus is sufficiently large, it allows the construction of prototypical intonation contours for each of the question types under investigation; this is one of the aims of the present study. Obviously, investigations should not rest at that. In a second step, the robustness of these prototypes has to be evaluated by mapping them onto questions in unpremeditated speech. However, given that a large-scale investigation of spontaneous questions was outside the scope of the present study, this part has to be left to future research[1]. Accordingly, the chapters 3 and 4 are based on a corpus of read speech, and the present chapter is expected to answer the following questions:

- Are the questions acoustically different from the statements in terms of (i) local f0 at the utterances' edges (i.e. onset height and final rise), and (ii) global f0 (i.e. overall trend and register level)?
- Do the differences follow the order predicted by the Functional Hypothesis, ST < WQ < YQ < DQ?

The first part of the chapter lays out the organisation of the production corpus. After some preliminary considerations in § 3.2, the hypotheses formulated at the end of the previous chapter are successively operationalised in § 3.3. Next, § 3.4 deals with the experimental design, the recording procedures, and the measurements and calculations carried out with respect to the chosen parameters and the independent variable Sex of Speaker. Following this, the second part of the chapter presents the results (§ 3.5) which, in § 3.6, are discussed in more detail. The chapter concludes with a general summary (§ 3.7).

[1] See chapter 2, § 2.4.

3.2 Preliminary considerations

3.2.1 Paradigmatic vs. syntagmatic approach

In the previous chapter it was hypothesised that Dutch questions, when compared with Dutch statements, are likely to feature

(i) a higher-pitched beginning
(ii) a final rise
(iii) a globally raised register
(iv) a raised nuclear accent peak
(v) less declination

As the intonation of statements is assumed to represent the default intonation from which question intonation may deviate, the questions in the corpus of read speech are compared with their (neutrally) assertive counterparts (although the lexical and/or syntactic make-up of the two cannot always be fully identical). Earlier, in chapter 1, three question types were singled out for investigation, the wh-question (WQ), the yes-no question (YQ) and the declarative question (DQ). Between these three, the hypothesised intonational properties were expected to be more salient in the absence of lexical and/or syntactic markers for interrogativity, i.e. inversion and wh-word (Functional Hypothesis).

Obviously, for the investigation of the parameters (ii) and (v) it suffices to make use of single utterances, as the relevant properties occur utterance-internally. Put differently, we expect possible differences to be primarily paradigmatic. Parameters (i), (iii) and (iv), however, may be given paradigmatic as well as syntagmatic operationalisations. In the former case, mean f0 values in the questions are compared with the corresponding mean f0 values in the statements. In the latter case, a given f0 value in an individual question is judged relative to the corresponding f0 value in an immediately preceding statement, on the idea that speakers may wish to use prosody to set off the specific discoursal function of a question against that of a statement. This requires the domain of investigation to be extended to sequences of two (pragmatically coherent) utterances, a move that has further advantages. First, it precludes variation in pitch level in the target utterances as might be prompted by the demands of discourse. Typically, the introduction of a new discourse topic causes initial pitch to be relatively high in the speaker's range (e.g., English: Brown *et al.* 1980, Beckman & Pierrehumbert 1986, Nakajima & Allen 1993; Dutch: Swerts, Bouwhuis & Collier 1994). We would not want this mechanism to interfere with our attempt to establish, for instance, whether questions start from higher pitch than do statements (cf. Hypothesis (i)). Making the target utterances members of a recurring lexical pair causes the novelty of the topic to wear off almost immediately (cf. Ladd & Terken 1995).

Moreover, embedding the target utterance in an utterance pair creates a context. There has been some dispute about the acceptability of citation forms in linguistic studies (cf. Brown *et al.* 1980:17, Oppenrieder 1988:173; Ladd 1996:198). Without the help of a context, subjects may be in the dark as to whether utterances are in broad or narrow focus. This may result in haphazard accenting or deaccenting of

constituents and introduce undesirable variability in the intonation contours. As it is assumed that a realisation under broad focus brings out unmarked accent placement and, therefore, the most neutral intonation contour, it seems desirable to guide the subjects towards broad focus by supplying an appropriate context. In addition, in order to increase the robustness of the results we decided to systematically vary the position of the target utterances within the pairs (first versus second position) and to combine them with context *statements* as well as with context *questions*. This yielded four logically different sequences, i.e. statement-statement, statement-question, question-statement, question-question. As for the context questions, only yes-no questions were included because sequences of, say, two declarative questions, or two wh-questions, proved pragmatically awkward. In § 3.3, the hypotheses (i), (ii), (iii) and (v) formulated above are operationalised; in the process, the experimental material is seen to gradually take its shape. To the testing of hypothesis (iv), a separate chapter is devoted (chapter 4).

3.2.2 Number of subjects

In production studies of intonation, there would seem to be no consensus as to what may be regarded an adequate number of speakers. Some studies would seem to implicitly subscribe to the idea that any speaker of a given language or dialect may be taken to represent his or her particular variety. However, as long as the extent of acceptable variation within that variety is unknown it seems hazardous to base the description of the intonation of (part of) a language on data produced by one or two speakers. Indeed, our pilot experiment demonstrated how such variation may mislead the investigator. The single female speaker in this pilot study produced question contours which had a number of fairly outspoken properties, such as falling initial boundary tones and a predominance of the low/rising nuclear pitch accent. Also, none of her wh-questions featured a final rise. Naturally, this led us to expect similar features in the questions spoken by the other nine subjects who, together with the speaker of the pilot experiment, later realised the full corpus of read speech. However, question contours produced by these nine were often substantially different. Clearly, there is some leeway, considering that the realisations of all ten speakers sounded canonical. What this suggested is that narrowing one's description to the performance of just one or two speakers holds a risk of underdifferentiation. To be sure, at present there is no principled way to decide on the truly appropriate number. But if the number of ten speakers (on whose data the present description is based) is as arbitrary as, say, five, or fifty, it is at least likely to capture a substantial part of the existing variation. And although heterogeneity in the data usually causes problems for statistical analyses, it seems preferable to a more homogeneous data set produced by a single speaker whose speech may in fact be relatively idiosyncratic.

3.3 Operationalisation of the hypotheses

3.3.1 Final rises

An utterance is assumed to have a final rise when pitch in the last (unstressed) syllable is auditorily perceived as going up and when this upward movement is reflected in the pitch curve. Presumably, the final rise occurs at two levels, depending on the level of the immediately preceding syllables. When these are phonologically low, the final rise is expected to take off from the baseline; when the preceding stretch is phonologically high, the final rise can be expected to take this high level for its starting point (cf. Pierrehumbert's 1980 'upstep rule').

3.3.2 Raised nuclear accent peak

Since the f0 level of an accent peak is best observed in relation to the height of an adjacent peak in the same utterance[2], target utterances are to be fitted with at least two (potential) accent slots. The issue of accent peak relations in, specifically, questions seems an interesting one, and comparing and contrasting statements and questions on this count is expected to provide useful information on possible prosodic differences between the two categories. Since a detailed investigation of the accent phenomena requires a separate chapter, the testing of hypothesis (iv) is postponed till chapter 4. Obviously, the scaling of successive accent peaks is closely bound up with focus structure. Although focus in statements has been the object of numerous studies, the issue of focus in questions seem to have attracted little attention. In the present study, an attempt is made to get some grip on this aspect of question intonation. Accordingly, the issue of focus in questions is, likewise, addressed in a separate chapter (chapter 6).

3.3.3 Raised onset

In order to get a broad impression of variation in onset levels in statements and questions, our first approach will be paradigmatic, that is, we compare the overall onset means of each of the four utterance types. In addition, it seems worthwhile to investigate how the f0 of an individual question onset relates to either onset or offset of an immediately preceding utterance. As is well known, there is a tendency in f0 to gradually decline across sequences of (statement) phrases or utterances (e.g. Swedish: Bruce 1982a; Danish: Thorsen 1985, 1986; Dutch (read speech): Sluijter & Terken 1993; Dutch (spontaneous monologue): Swerts & Geluykens 1994). This phenomenon, sometimes termed 'supradeclination', should be kept distinct from downtrends which can be observed utterance-internally, viz. declination, downstep and final lowering (cf. chapter 2, § 2.2.2.4.2). Within a paragraph, the ongoing downtrend is commonly checked by a reset causing a next clause or utterance to

[2] Obviously, it is also possible to look at peak differences from a paradigmatic rather than a syntagmatic angle, i.e., by comparing the mean f0 values of all second accent peaks to the mean f0 values of all first accent peaks. Nonetheless, we prefer the syntagmatic approach as it may shed light on speakers' strategies for expressing certain utterance-*internal* relations between subsequent peaks (cf. Liberman & Pierrehumbert 1984).

begin at higher pitch. Clearly, this reset between clauses or utterances is a syntagmatic phenomenon. So far, however, it has been mainly investigated in sequences of *statements*. Now, in *questions* an increased reset may well act as a prosodic device to alert the listener to the interrogative status of the utterance. After all, unlike statements questions require a listener to respond. Hence, when it is the case that questions have raised onsets, we expect reset between a statement and a *question* to be greater than between a statement and a *statement*.

To assess the amount by which f0 is reset, various methods have been used. Some researchers have computed reset by comparing the peak heights of consecutive utterances. Others have looked at f0 values of hand-fitted declination lines (toplines, baselines or both), comparing the onset value of utterance II with the offset, or onset, value of utterance I (e.g. Cooper & Sorensen 1981; Ladd 1984, 1988; Thorsen 1985, 1986). As regards reset in Dutch, Collier (1987) investigated declination reset at syntactic boundaries between main clauses and subclauses of complex sentences. Being interested in two possible manifestations of reset, he compared declination baseline onsets and offsets, investigating (i) whether the **on**set of the second clause was higher than the **off**set of the first, and (ii) whether the **on**set of the second clause was higher than the **on**set of the first[3]. As regards (i), which might be termed 'local reset', an average upward shift of 1.7 semitone (ST) was found. However, for (ii), which might be called 'global reset', there was no systematic evidence. This is not really surprising, considering that the experiment was concerned with sentence-*internal* onsets and offsets. Within a (complex) sentence, the prosodic cohesion between the constituent parts is still fairly strong (cf. Thorsen 1985, 1986). An experiment carried out by Sluijter & Terken (1993) explicitly compared onsets[4] *across* sentences occupying initial, medial or final positions in a paragraph. Both topline and baseline onsets failed to display global reset, that is, the onsets of utterances in medial or final position were lower than in initial or medial position, respectively.

The approach to reset taken in the present study slightly deviates from that in the above studies. Unlike Collier and Sluijter & Terken, we will not compare (extrapolated) onset values of declination lines, since we intend to measure the global f0 trend by a different method (see § 3.3.4 below). Instead, it is raw onset values that will be compared. Our data are unlike those of Collier in yet another respect, i.e., speakers do not receive instructions regarding the phonological form of onset or offset (i.e. low vs. high boundary tone). Indeed, their choice of boundary tones in questions is one of the objects of study. This implies that the raw onset values are not split up along the phonological dimension %H vs. %L[5].

With the above considerations in mind we arrive at the following operational definition of the parameter ONSET HEIGHT. First, the amount of *local reset* is defined as the difference between the offset value of utterance A and the

[3] The set of sentences comprised complex statements as well as complex questions. Comparisons were made across onsets/offsets that were phonologically equal, i.e. low-low, or high-high, with the latter providing information on topline resetting. Of this, no evidence was found.
[4] Extrapolated from the declination toplines and baselines.
[5] In anticipation of the results of the transcription reported in chapter 4 it may be mentioned here that, in the questions, 89% of the initial boundary tones were transcribed low (%L) although, phonetically, they were found to be higher than in the statements.

onset value of utterance B, as illustrated in figure 3.1a. When local reset occurs, it is hypothesised to be larger between statements and subsequent *questions* than between statements and subsequent *statements*.

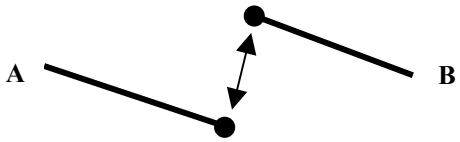


Figure 3.1a. Local reset (offset~onset).
The arrow indicates the amount of reset between the **off**set of utterance A and the **on**set of utterance B.

Second, the amount of *global reset* is defined as the difference between the onset of utterance A and the onset of utterance B. Again, when global reset occurs, we expect it to be larger in *questions* following statements than in *statements* following statements. When the second onset is lower than the first, we call it 'incomplete reset'. For two equal onsets, we use Collier's term 'complete reset'. When the second onset is higher, this is termed 'overcomplete reset' (cf. Figure 3.1b).

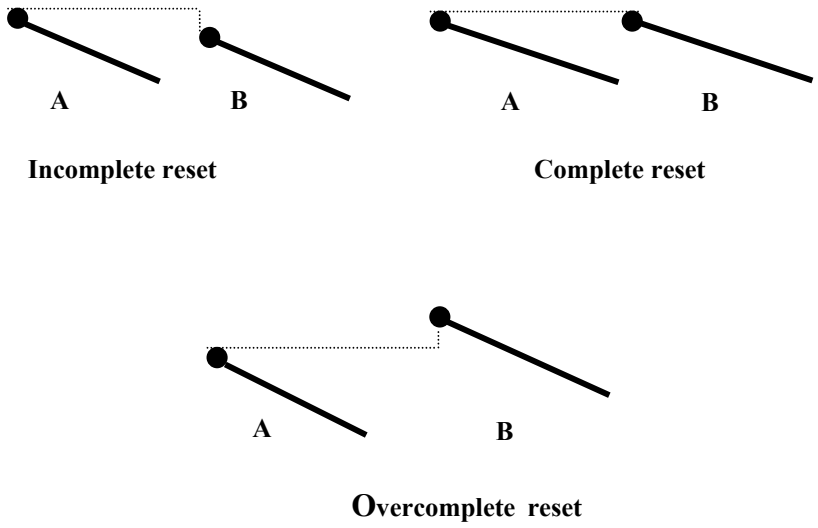


Figure 3.1b. Global reset (onset~onset; see text)

3.3.4 Raised register level

Not only are questions hypothesised to have higher onsets than statements (a local phenomenon), their global pitch level is also predicted to be higher, i.e., questions are

expected to be realised on a raised register level. In chapter 2, register was defined as the span between the lowest and the highest tone within a given utterance ('register span'). While this definition primarily refers to register *width*, the lowest f0 value is also indicative of the register's position within the overall range of a speaker, i.e. of register *level*. An upward shift of this level is assumed to be reflected in raised f0 minima (cf., e.g., Swerts *et al*. 1994:2065).

3.3.5 Global trend

Establishing whether interrogativity causes the utterance-internal downtrend ('declination') to be reduced, or even reversed into inclination, is not an easy matter since declination is notoriously hard to quantify (cf. Vaissière 1983:56). It may be recalled here that, at an early stage of the present research, a pilot production experiment was carried out in order to get some first impressions of the intonation of Dutch questions. Twenty pairs of questions and statements were produced by a female speaker and traditional declination toplines and baselines were fitted to the f0 peaks and valleys by eye ('visual abstraction method'). On the whole, the results seemed to confirm our expectations as laid down in the hypotheses above. Nonetheless, this exploratory experiment led us to abandon the visual abstraction method for the representation of overall f0 trends. For although this method has been widely used (e.g., 't Hart *et al.* 1990, Sluijter & Terken 1993), Lieberman, Katz, Jongman, Zimmerman & Miller (1985) have pointed out that it is essentially subjective and difficult to reproduce. As a more reliable means of describing the global direction of an intonation contour, the authors propose an all-points linear regression line which can be easily and objectively calculated by computer[6]. However, calculation of an all-points line does not produce top lines and base lines, which are obviously indispensible for, e.g., assessing register span. Therefore, it was decided to extend Lieberman *et al.*'s proposal in the following manner:

(i) For each utterance, an *all-points* linear regression line was calculated for f0 as a function of time. This all-points line was not meant to have a status of its own; it merely functioned as a tool for dividing the intonation contour in an upper and a lower half.
(ii) A *lower trend line* was then fitted through the f0 points located in the lower half.
(iii) Similarly, an *upper trend line* was fitted through the f0 points located in the upper half.

It is important to note that the predictions of a regression model differ somewhat from those of a declination model. While the latter encompasses the full register span defined by the distance between a high declination line (intersecting with the local peaks) and a low declination line (intersecting with the local valleys), the area demarcated by the regression lines will be narrower by roughly 50% because

[6] For a discussion of pros and cons of this method, see the exchanges in Repp (1985), JASA 78 (3), 1114-1115, Lieberman *et al*. (1985), JASA 78 (3), 1116-1117, 't Hart (1986), JASA 80 (6), 1838-1839, Lieberman (1986) JASA 80 (6), 1840-1841.

regression lines, by definition, gravitate away from extreme values. As a result, regression slopes will generally be less steep. By the same token, a regression model will have a levelling effect on variation in register span, given that such variation largely manifests itself in the outermost areas, i.e. in the peaks and valleys. In a regression model, these areas fall outside the area enclosed by the regression lines. In effect, in a regression model pitch variation affects slope and register span only indirectly, whereas in a declination model this link is direct. In the interpretation of the (statistical) results, this overall levelling effect of the regression model should be taken into account.

In connection with the regression model, there is yet another point that requires some attention. *A priori*, statements and questions were assumed to differ in that the latter would frequently have final rises. Such rises were bound to give an upward slant to all-points regression lines, making it impossible for overall trends of statements and questions to be fruitfully compared[7]. To obviate this, the data points of the final rises were excluded from the calculation of the regression lines.

Figure 3.2 illustrates the procedure of fitting the regression lines.

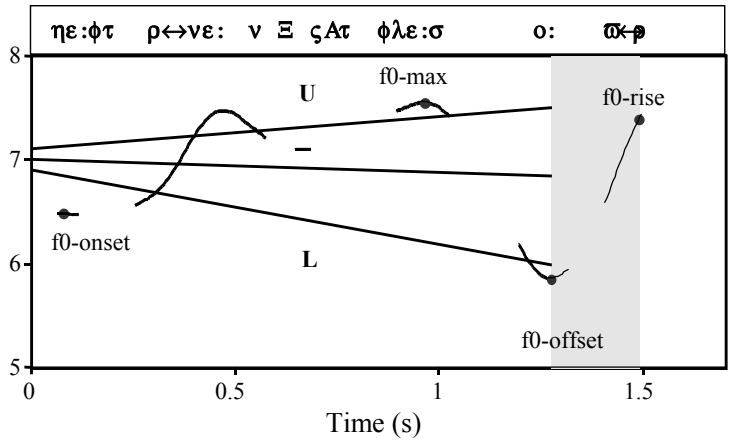


Figure 3.2. Fitting of upper (U) and lower (L) regression lines to the yes-no question *Heeft Renée nog wat vlees over?*

[7] As a matter of fact, interference from the final rise hampered Eady and Cooper's (1986) comparisons of the maximum f0 of sentence-final words in statements and their question versions. Considering that three out of the four experimental sentences featured an accent on the final syllable ('pool', 'school' and 'Japan', as opposed to 'Sunday'), it is very likely for maximum f0 to have also included the final rise. In the more recent AM-framework (cf. Ladd 1996), this terminal portion is analysed as consisting of a sequence of a pitch accent (T*) and a high boundary tone (H%).

3.3.6 Sex of Speaker

3.3.6.1 Sex-related differences in speech

In chapter 2 (§ 2.2.5), it was argued that female speech is likely to have a wider register span[8] than male speech, given empirical evidence that it is 'more varied' and 'more expressive'. In acoustic terms, these two properties may be expected to show up as greater variation around the average pitch, that is, as a wider register span. Nonetheless, a comparative survey of 17 studies on register span in males and females reported the opposite (Henton 1989). In order to ascertain whether the alleged 'swoopiness' of female speech is reality or, rather, a stereotyped idea, Henton transformed the results of these 17 studies from linear Hertz values into logarithmic semitones (ST). This enabled her to compare the male and female ranges in terms of perceptually realistic units and to conclude that "it is the males who generally have a greater pitch range than females have".

However, some caution is called for, as Henton herself points out. First, the 17 studies were fairly heterogeneous in that they comprised semi-spontaneous as well as read speech, and isolated vowels as well as full stretches of text. Besides, the subjects were speakers of different languages/varieties, i.e. American English, British English, Polish and Swedish. Also, in only five out of the 17 studies the data had been simultaneously collected from both sexes; the other studies were concerned with either male range or female range. In the majority of the studies, there was no guarantee that the subjects had been selected with a view to homogeneity in age, educational background, etc. Finally, it was not always clear whether the results had been based on the entire pitch range or on the middle 90% or 95%. Obviously, these uncertainties and imbalances in the body of 17 studies weakened the conclusion that the male register was generally the wider of the two.

With these reservations in mind, Henton carried out an experiment of her own. She compared the pitch ranges of a homogeneous group of five North American men and five comparable women who read out 90 s. of speech. Taking the central 95.4% of the speakers' f0 distributions into account, she found a slightly wider register (in ST) for males in one condition, and the reverse in the other condition[9]; differences were not significant. This led her to relegate the claim that female speech is 'swoopy' in terms of register width to stereotype.

As far as Dutch is concerned, a number of studies have compared female and male pitch ranges (i.e. register span) in spontaneous, semi-spontaneous and read stretches of speech. In none of these studies were significant sex-related differences in register found. In Tielen (1992), read and spontaneous utterances of ten men and ten women were investigated (in ST). In Kraayeveld (1997), the same was done with

[8] Often, the literature uses the term 'pitch range' for what the present study refers to as 'register width' or as 'register span' (cf. Ladd 1996:260). For the sake of consistency, however, we stick to Clements' (1990:59) definitions of range and register given in chapter 2 and repeated here: "*Range* refers to the largest interval (or frequency band) within which tones are normally produced in the speech of a given speaker. *Register* is a smaller interval or frequency band internal to the speaker's range, which determines the highest and lowest frequency within which tones can be realized at any given point in an utterance".

[9] These conditions, involving the presence vs. absence of nasal segments, were not relevant for pitch range.

respect to read and semi-spontaneous speech of 25 men and 25 women (in ST). In the spontaneous part of the material, females displayed a wider register than did males. However, in the read speech (i.e. a passage, and sentences giving sports results) the reverse was found. Neither were there significant differences in the excursion sizes of the pitch accents. Biemans (1998) compared spontaneous conversations across the sexes (57 men, 57 women, in ST), and found no significant differences. Similarly, in Van Donzel *et al.* (1998) the differences in range (in ST) between four men and four women in semi-spontaneous speech fell short of significance.

These results on Dutch seem to suggest that Henton was right in concluding that differences between male and female speech as posited in the literature do not correspond with differences in register span.

3.3.6.2 Sex-related communicative behaviour

However, there would seem to be two reasons for believing that such a conclusion may be premature. First, the studies mentioned above predominantly dealt with *declarative* speech and the question arises whether the results would have been the same in *interrogative* speech. There is empirical evidence that the communication patterns of men and women have different characteristics. Men have been reported to be more direct and assertive, as well as to prefer making statements and carrying out tasks; on the whole, their communicative behaviour aims at 'agency'. Women, by comparison, are claimed to be more inclined to strive for 'communion'; their personal involvement is greater and this is reflected in more emotional expression and a greater inclination to listen (e.g. Meeuwissen 1988, Basow 1992). Besides, men and women are reported to adopt different strategies to influence other people. Women would seem to be less afraid than men to ask for help or to reveal lacunae in their state of knowledge or information (cf. Meeuwissen 1988). In all, female communicative behaviour is generally associated with a greater preparedness to show involvement as well as dependence; primarily, it is directed towards interaction.

Obviously, the communicative function interrogativity is ideally suited for establishing interaction. Given the above claims about sex-related differences in communicative behaviour, one might expect these to find intonational expression primarily in interrogative speech, more so than in declarative speech. That is, not only can women be expected to be more inclined to ask questions than men, they are also more likely to make maximal use of the intonational devices that go with questions. In chapter 2 we saw that, across languages, questions are typically higher-pitched than statements. If women are less afraid to show the dependence implicit in questioning, we hypothesize that their questions will display a relatively large quantity of high pitch[10].

There is a second reason for not yet accepting Henton's rejection of sex-related differences in register span. Not only did the studies carried out so far largely concentrate on *declarative* speech, they mainly investigated *global* span. Our

[10] With respect to the parameter DOWNTREND, 'higher pitch' means 'a less downward overall trend'.

approach means to break up utterances into *local* pitch configurations (i.e. pitch accents and final boundary tones), the span of each of which is then compared across the sexes. Specifically, we expect the local pitch accents in the questions as well as the boundary tones reflecting final rises to have larger excursions in the women than in the men. To our knowledge, sex-based differences in final rises have so far only been dealt with in a study of statements and (declarative) questions in Canadian-French (Ryalls *et al.* 1994). Here, the difference between statements and questions was defined as the difference between the average f0 of the last syllable of a question minus the average f0 of the last syllable in the matching statement. After normalisation for the intrinsic difference in average f0 between the sexes, values on this parameter were still found to be significantly higher in the female than in the male speakers.

Obviously, it cannot be denied that global measures are related to local measures. Nevertheless, by differentiating between the two we hope to pinpoint possible local differences between the sexes that fail to show up in global measures.

3.4 Method

3.4.1 Material

The design for the production corpus was based on two core utterances, each with two accentable syllables containing identical vowels so as to preclude interference by intrinsic pitch: *Renée heeft nog vlees over* ('Renée has still meat left'), and *Marina wil haar mandoline verkopen*, ('Marina wants her mandolin sell'). While the former utterance consisted of seven syllables (with two unstressed syllables between the first and the second accent, and two following the nuclear accent), the latter had twelve syllables (with five unstressed syllables between accents, and four following the nuclear accent). This division in 'short' and 'long' utterances was introduced to compare and contrast the behaviour of f0 in stretches of unaccented speech, e.g. with a view to declination and the timing of the final rise. For instance, it has been claimed that long utterances have less steep declination slopes than short utterances (cf. 't Hart *et al.* 1990:128). Considering that questions are expected to display less declination than statements (or none at all), predictions are that longer questions will have even shallower slopes than short questions.

In their basic forms, the two core utterances served as statements and as declarative questions. With inversion, they were transformed into yes-no questions; with inversion and an additional wh-word, into wh-questions. Each of the two core utterances was paired with a suitable context sentence so as to form minimal paragraphs (*Onze poes moet wat eten hebben*, 'Our cat must some food have', and *Er is donderdag weer een rommelmarkt*, 'There is thursday again a jumble sale', respectively). With a view to making the design orthogonal, we also included pairs consisting of two statements or two questions. However, as pairs of two consecutive wh-questions or two declarative questions sounded awkward, we used only yes/no questions for context questions. The positions of target and context utterances were systematically varied, such that the eight target utterances could occur in first or in second position. In addition, the target utterances were also realised in isolation. This resulted in 32

utterance pairs and 8 isolated utterances, i.e. in 40 stimuli. These stimuli were semi-randomised (i.e., no immediate successions of the same basic sentence were allowed) and presented to speakers on separate cards. The Tables 3.1a and 3.1b present an overview of the experimental material.

Table 3.1a. Overview of utterance pairs.

Target utterance (short)	Context utterance (short)
ST Renée heeft nog vlees over. /rəne: he:ft n Ξ fle:s o:vər/ **YQ** Heeft Renée nog wat vlees over? /he:ft rəne: n Ξ ɕAt fle:s o:vər/ **WQ** Wat heeft Renée nog voor vlees over? /ɕAt he:ft rəne: n Ξ fo:r vle:s o:vər/ **DQ** Renée heeft nog vlees over? /rəne: he:ft n Ξ fle:s o:vər/	**ST** Onze poes moet wat eten hebben. /ɔnzə pus mut ɕAt e:tə hEbə/ **YQ** Wil de poes nog wat eten hebben? /ɕIl de pus nɔΞ ɕAt e:tə hEbə/
Target utterance (long)	**Context utterance (long)**
ST Marina wil haar mandoline verkopen. /ma:rina: ɕIl ha:r mAndo:linə vərko:pə/ **YQ** Wil Marina haar mandoline verkopen? /ɕIl ma:rina: ha:r mAndo:linə vərko:pə/ **WQ** Waar wil Marina haar mandoline verkopen? /ɕa:r ɕIlma:rina: ha:r mAndo:linə vərko:pə/ **DQ** Marina wil haar mandoline verkopen? /ma:rina: ɕIl ha:r mAndo:linə ɷərko:pə/	**ST** Er is donderdag weer een rommelmarkt. /Er Is d ndərdAΞ ɕe:r ən r məlmArkt/ **YQ** Is er donderdag weer een rommelmarkt? /iz ər d ndərdAΞ ɕe:r ən r məlmArkt/

Table 3.1b. Schematic representation of the design. Each cell frequency must be multiplied by four (short/long utterance, two repetitions).

Type of target utterance	Context utterance				Isolation	Total
	ST		YQ			
	1st pos	2nd pos	1st pos	2nd pos		
ST	10	10	10	10	10	50
WQ	10	10	10	10	10	50
YQ	10	10	10	10	10	50
DQ	10	10	10	10	10	50
Total	40	40	40	40	40	200

3.4.2 Speakers and recording procedures

Subjects were ten native speakers of Standard Dutch between 20 and 48 years old (five male, five female), all of them students or university staff members. They twice read out the set of 40 stimuli, half of them in a reverse order so as to counterbalance possible order effects. Subjects were not aware of the goals of the experiment. They were instructed to read out the material in a lively and realistic manner, as if they were actors in a radio play; this was expected to yield a certain range of prototypical statement and question contours. When a speaker made an error he or she was asked to repeat the sentence (pair). As regards accentuation, no instructions were given. Recordings were made on digital audiotape (48.1 kHz, 16 bits) in a soundproofed studio. They resulted in a corpus of 800 target utterances, i.e. 200 statements, 200 wh-questions, 200 yes-no questions and 200 declarative questions, as well as 640 context utterances.

3.4.3 Measurements

F0 was extracted by the method of subharmonic summation (Hermes 1988), followed by curve-smoothing over a 50-ms time window. F0 values were expressed in ERB (Equivalent Rectangular Bandwidths), which is currently held to be the psychophysically most relevant scale in intonation research (cf. Hermes & Van Gestel 1991, Ladd & Terken 1995, Terken & Hermes 2000)[11]. The data were processed with the speech processing package PRAAT (Boersma 1998). The presence of a terminal rise was determined interactively in the visual display of the f0 curve. The procedure of

[11] Formula: E=21.4*log10(0.00437*Hz+1)

fitting upper and lower trend lines as described above (§ 3.3.4) was carried out. Crucially, these regression lines were fitted to the data points *minus* the final rise (if present), in order to make statements and questions mutually comparable. Figure 3.3 shows the individual data points whose f0 values were measured:

1. Utterance onset f0, i.e. f0 of the first voiced frame (or: classified as voiced by the pitch extraction algorithm);
2. Minimum f0;
3. Postnuclear low (PL), i.e. in statements (and in questions without a final rise), the last f0 value of the utterance; in questions, the f0 value of the onset of the final rise (when present);
4. Highest f0 of the final rise (if present).

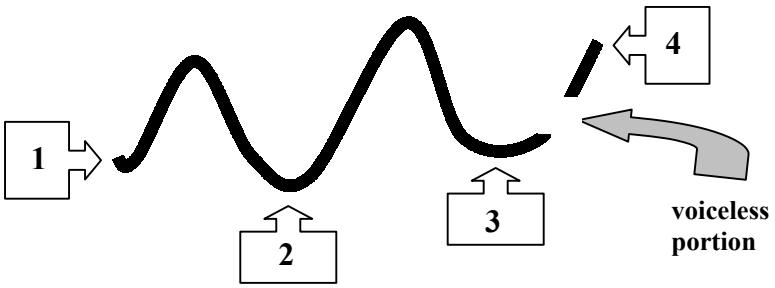


Figure 3.3. Measuring points.
(1) onset f0, (2) minimum f0, (3) postnuclear low (PL), and (4) highest f0 of final rise.

The final rise always occurred in the last syllable (which was unaccented and usually included a voiceless portion[12]). Its excursion size was defined as the interval between data point 3 and 4. With a view to the calculation of reset, the utterance onset f0 and utterance offset f0 of the context sentences were also measured.

With respect to the term 'postnuclear low', some explanation is in order. The problem with data point 3 is that its status in a statement differs from that in a question, given that the latter usually features a final rise. Thus, whereas in a statement point 3 represents the utterance's *offset* (which presumably includes final lowering), in a question it represents the *onset* of the final rise. To prevent terminological confusion, we decided to use the more neutral term 'postnuclear low' (PL) for point 3. Nonetheless, in the sections which explicitly deal with the final rise it is usually more insightful to refer to the rise's starting point as the '(rise) onset' and to its end point as the '(rise) offset'; we trust that this usage will be clear from the context. Note that Figure 3.3 does not include more specific f0 measurements of the two pitch accents;

[12] In the short core utterance, the onset of the final (unaccented) syllable was voiceless /p/ (*verkopen*); in the long core utterance, this onset was frequently devoiced (*over*).

this part of the investigation is reported in chapter 4.

3.5 Results

The data were submitted to a number of separate analyses of variance (SPSS Manova, repeated measures) with one fixed between-subject factor: SEX OF SPEAKER (two levels, male and female), and three fixed within-subject factors: UTTERANCE TYPE[13] (four levels: statement, wh-question, yes-no question, declarative question), UTTERANCE LENGTH (two levels: short and long), and UTTERANCE POSITION (three levels: in isolation, in first, and in second position within a pair). SUBJECT was nested within SEX OF SPEAKER as a random factor. Figure 3.4 illustrates the basic design of the analyses of variance.

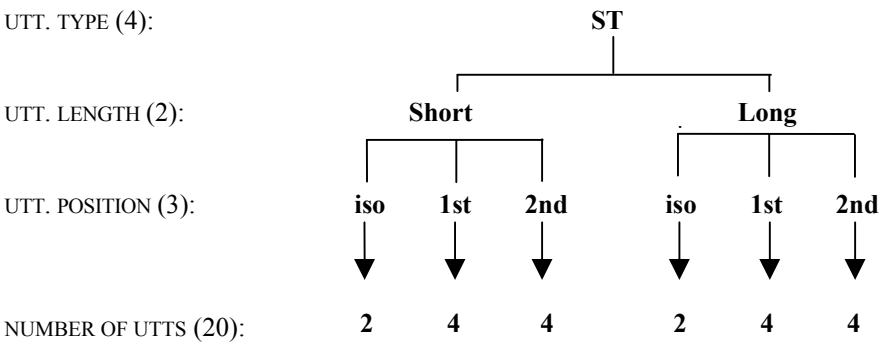


Figure 3.4. Design of SPSS Manova (repeated measures), exemplified for utterance type ST. For the other three utterance types (WQ, YQ and DQ), the structure was the same.

Since we had specific predictions for each of the dependent variables, we preferred separate analyses to a single analysis including all variables. A further argument for separate analyses was that sometimes part of the data had to be excluded, depending on the hypothesis. To give an example, final rises could start from a level that was either phonologically low ('low H%') or phonologically high ('high H%')[14]. To lump these two types together in a calculation of, say, the mean onset f0 of final rises would be to misrepresent reality. Thus, either one or the other category had to be excluded from the analysis. As a consequence, it was possible for cells to remain empty. Such cases of missing values would have rendered a simultaneous analysis of all dependent variables impossible, given that a single missing (mean) value

[13] Whenever this factor applies to the questions only, we use the term QUESTION TYPE rather than UTTERANCE TYPE.

[14] For diagrams of these configurations, see chapter 4, Figures 4.1a and 4.1b.

causes the SPSS statistical programme to discard the entire case (i.e. the speaker) from which the value is missing. Obviously, the possible exclusion of subjects from the analysis was highly undesirable. To be sure, this practical problem also cropped up in the individual analyses, but here it could be individually remedied by applying the procedures outlined below.

In most analyses, 24 different values had been entered for each of the ten subjects (i.e. 4 utterance types x 2 lengths x 3 positions, cf. figure 3.4 above). Each value represented either the means of two utterances (i.e. of those presented in isolation), or the means of four utterances (i.e. of those presented in pairs). Cells could be empty for several reasons. For instance, one of the subjects never realised final rises on her wh-questions. Accordingly, her case displayed empty cells in all three wh-conditions, that is, in the isolated wh-questions as well as in the wh-questions occurring in first and second positions in the utterance pairs. To give a second example, while subjects might have produced one or more final rises in the (4) paired occurrences of the wh-question, thereby producing a (mean) value to fill the cell, final rises on the (2) wh-questions occurring in isolation might have been lacking. This would cause the corresponding cell to be empty, leading to the loss of all other values pertaining to the subject in question. To prevent this, we filled empty cells with estimates derived from the available data. When the empty cell occurred in one out of the three positions of the target utterance (e.g., in wh-questions in isolation), we filled it with the mean value of the same subject on the other two positions (e.g., with the mean of the wh-questions in first and second position). When, however, all three positions yielded empty cells (as was bound to happen in the person who never realised a final rise in wh-questions), we filled it with the overall mean value of this subject on the parameter in question (e.g., with the mean size of all final rises realised in the other conditions). Note that this particular problem arose mainly in connection with the parameter FINAL RISE; other parameters such as ONSET HEIGHT or OVERALL TREND did not cause difficulties because there were no missing values. Besides, while the missing data treatment is implemented in the Manova's, the graphs and tables below present the actual means.

Since a large number of statistical analyses increases the risk of wrongly rejecting the null-hypothesis, the level of significance was fixed at .01 rather than .05. In cases of significance, we report Huyn-Feldt corrected significance values. Interactions are mentioned only when significant. For post hoc analyses, use was made of Tukey tests; here, a significance level of .05 was felt to be acceptable. As *p* values were not always comparable due to differences in degrees of freedom, we also report correlation ratios (η^2) indicating the effect strength (cf. Rietveld & Van Hout 1993:59). It must be noted, however, that multiplication of the value for η^2 by 100 does not yield the contribution to SS_{tot}, as is the case when η^2 represents SS_{effect} /SS_{tot}, causing the sum of η^2 effects to add up to 1. Rather, the value for η^2 as supplied by the SPSS statistical package represents a partial effect, i.e. SS_{effect}/ SS_{effect}+ SS_{error}, and the sum of these η^2 effects does not by definition add up to 1.

Given that the anatomy of women causes the frequency of female speech to be roughly twice as high as that of male speech, it is only natural for the factor SEX OF SPEAKER to have a significant effect on all dependent variables expressed in raw f0 values. Hence, it makes little sense to report these main effects. Instead, raw values will

only be considered *within* the male or the female frequency domain. Unless stated otherwise, results of post hoc tests are generally collapsed over the sexes. At times, however, it proves more insightful or more relevant to break down results or figures by SEX OF SPEAKER, or even by individual subject. For parameterised variables such as final rises it is obviously perfectly possible to compare male and female performances; for this goal, ERB is taken to be an appropriate scale (cf. Ladd & Terken 1995).

3.5.1 Final rises

3.5.1.1 Incidence

As expected, in the 200 statements (ST) not a single final rise occurred. By contrast, in the 200 declarative questions (DQ) and the 200 yes-no questions (YQ) final rises were produced in 200 and 192 cases, respectively. The number of final rises in the 200 wh-questions (WQ) amounted to 128, which is quite high in view of crosslinguistic claims that wh-questions go without final rises (cf. Hirst & Di Cristo 1998). Figure 3.4a shows the performance across the ten speakers.

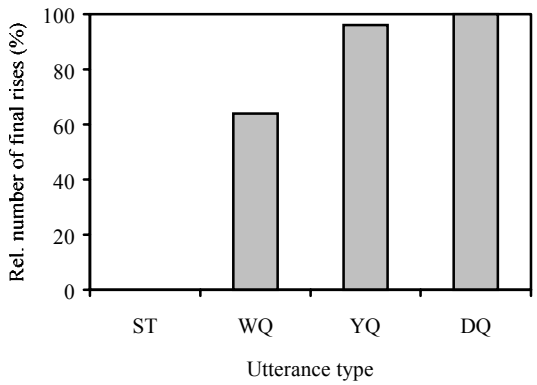


Figure 3.4a. Final rises (across all speakers).
Percentages of final rises in the statements (ST), wh-questions (WQ), yes-no questions (YQ), and declarative questions (DQ).

Between the ten subjects, however, there was some variation. Three speakers produced final rises in all three question types in all conditions. With the others, notably the wh-questions frequently lacked a final rise, as Figure 3.4b shows. Typically, speaker F5 never ended her wh-questions with a final rise. She had been the sole speaker in the pilot experiment mentioned in § 3.2.2.5 and her earlier performance had incorrectly biased our expectations as regards the incidence of the final rise in this question type.

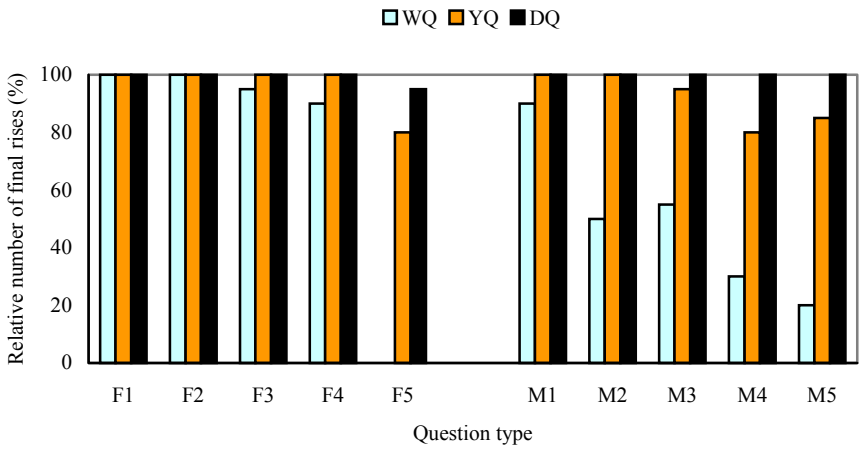


Figure 3.4b. Final rises (by speaker)
Percentages of final rises in the question types WQ, YQ, and DQ broken down by individual speaker (F1-5 are female, M1-5 are male).

3.5.1.2 Excursion sizes[15]

An initial observation was that the three question types featured roughly equal-sized final rises: there was no significant effect of the factor UTTERANCE TYPE. Individual Anovas for the ten speakers confirmed this: in only one speaker did the excursion differences between the question types exceed the .01 level of significance[16]. If anything, on average the rises in wh-questions were the largest, in declarative questions the smallest; this ran counter to the Functional Hypothesis. Figure 3.5 presents the respective mean sizes pooled across the ten speakers.

[15] In a small number of cases, sizes could not be computed as the f0 of rise onset or rise offset could not be properly defined due to creak or whisper.
[16] This applied when high H% and low H% were lumped together, as well as when the subset of the (less common) high H% was excluded; see further below.

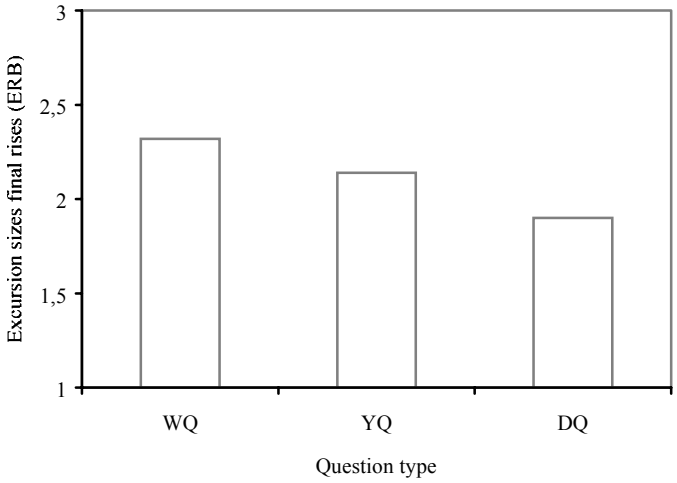


Figure 3.5. Final rises (excursion sizes)
Mean excursion sizes (ERB) of final rises in the four utterance types collapsed over the ten speakers.

However, closer observation suggested that reality was more subtle. First, across the question types, the f0 levels from which the final rises took off[17] stepped up following the hypothesised order WQ<YQ<DQ. This difference was significant: $F_{(3,24)}=16.27$, $p<.001$, $\eta^2=.670$. Post hoc comparisons indicated that the statements differed from each of the three question types; also, WQ differed significantly from DQ ($p=.004$, see further § 3.6.4 below). At the same time, the rises' *final* frequencies also successively stepped up, in the same order. While this accounted for the relatively fixed excursions of the final rises across the question types, it also meant that these rises were typically realised in increasingly higher sections of the speaker's range (i.e. WQ<YQ<DQ). Figure 3.6 illustrates this for both sexes. Clearly, this stepwise increase of the pitch levels made up for the lack of distinction between the rises' magnitudes.

[17] I.e., the Postnuclear Lows (PL), see Figure 3.3 above.

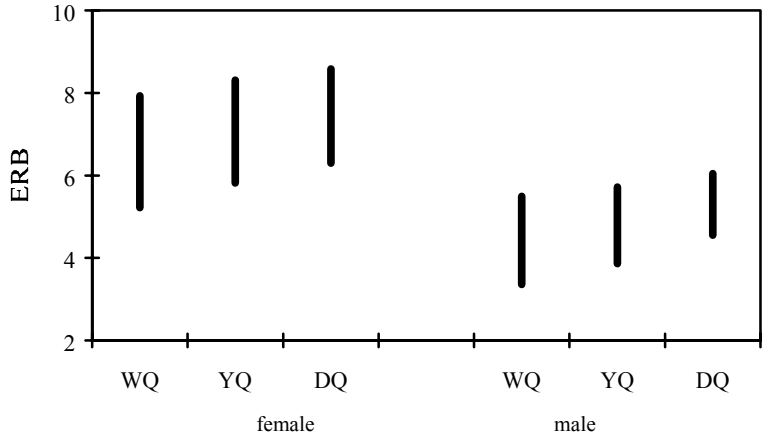


Figure 3.6. Final rises: interpolations between mean rise onset values (PL's) and mean rise offset values (in ERB), broken down by QUESTION TYPE and SEX OF SPEAKER.

Next, we differentiated between the two *phonologically* determined levels of rise offset, that is, between the final rise that set off from the baseline ('low H%') and the final rise that set off from a phonologically high level ('high H%'). It emerged that excursions of low H% were considerably larger than those of high H%, as shown in Figure 3.7. This is only to be expected, considering that the starting point of high H% is closer to the ceiling of a speaker's range.

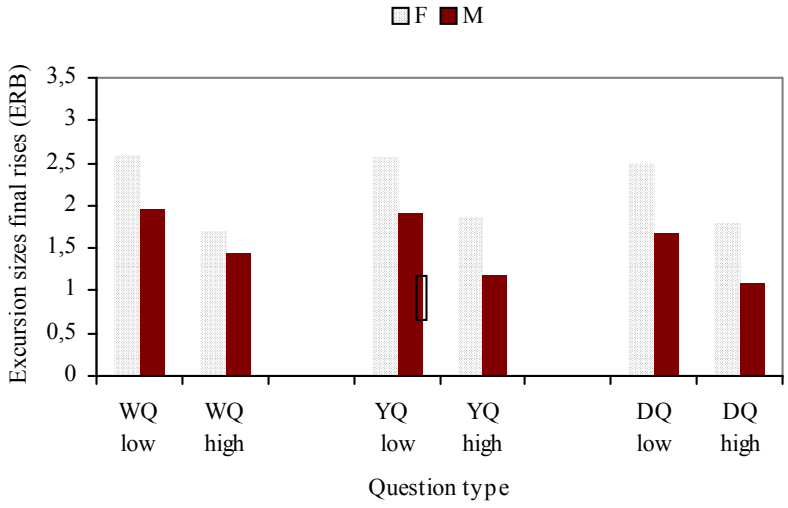


Figure 3.7. Final rises ('low H%' vs. 'high H%')
Mean sizes (ERB) of final rises broken down by QUESTION TYPE, SEX OF SPEAKER and type of rise.

Third, the final rises produced by the women were markedly larger than those produced by the men, as Figure 3.7 above makes also clear. Statistically, however, the effect of the factor SEX OF SPEAKER fell just short of the pre-set significance level of .01 ($F_{(1, 8)}$=7.53, p<.025, η^2=.485). Table 3.2 gives the mean sizes for each question type for the male and the female speakers, together with the differences between the sexes (last column). Note that, *within* the sexes, sizes were fairly constant, as were the differences *across* the sexes. The sex-related aspects of the results will be further discussed in § 3.5.5 below.

Table 3.2. Final rises (excursion sizes)
Mean excursion sizes (ERB) of final rises in wh-questions (WQ), yes-no questions (YQ) and declarative questions (DQ), broken down by SEX OF SPEAKER. Between brackets: standard deviations.

Question type	Female	Male	Δ F-M
WQ	2.58 (.70)	1.92 (.77)	+ .66
YQ	2.45 (.66)	1.85 (.63)	+ .60
DQ	2.29 (.83)	1.50 (.64)	+ .79

It is only fair to add, however, that the performance of the individual speakers as given in Figure 3.8 below somewhat detracts from the apparent robustness of the excursion size of the final rise. Also, there is some overlap across the sexes: the excursions of three female speakers (F1, F2, F5) are more or less equal to those of two male speakers (M3, M5); the overall difference between the sexes appears to lie in the performance of the remaining two female and three male subjects.

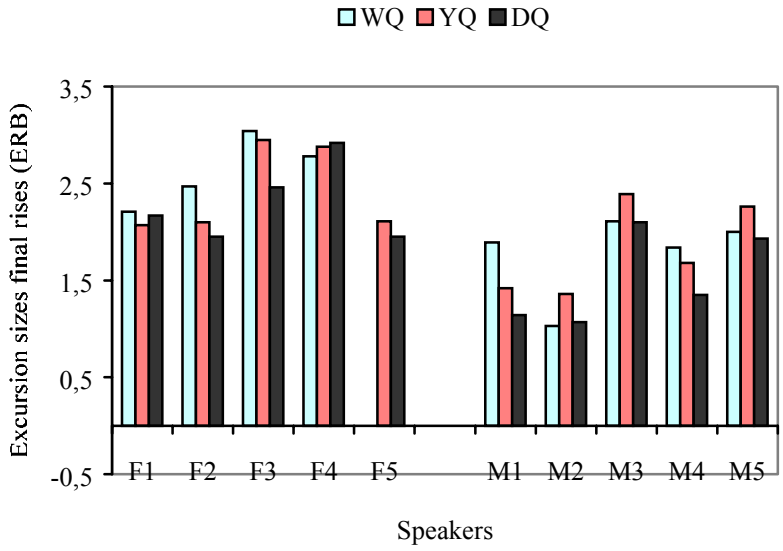


Figure 3.8. Final rises
Mean excursions of final rises (ERB) broken down by individual speaker (F is female, M is male).

The factors UTTERANCE LENGTH and UTTERANCE POSITION did not show significant main effects on the excursion sizes of the final rise. One significant interaction was found: SEX OF SPEAKER x QUESTION TYPE x UTTERANCE LENGTH ($F_{(1, 16)}$=7.16, p< .006). That is, in the female speakers, the short versions of WQ and YQ had larger final rises than the long ones; in DQ, there was hardly any difference between the long and short versions.

When we sum up the results for the excursion sizes of the final rises, the following picture emerges:

(i) excursion sizes did not significantly differ as a function of QUESTION TYPE; however:

(ii) onsets of the final rises were successively higher, in the order hypothesised by the Functional Hypothesis (WQ<YQ<DQ);

(iii) offsets of the final rises were successively higher, likewise following the order hypothesised by the Functional Hypothesis;

(iv) final rises taking off from the base line ('low H%') were larger than final rises taking off from a phonologically high level ('high H%');

(v) final rises were larger in the female than in the male speakers, notably in the short questions;

3.5.2 Utterance onset height

3.5.2.1 Paradigmatic approach

First, this variable was approached from the *paradigmatic* angle. That is, we determined whether, overall, the mean values of the utterance onsets were higher in the questions than in the statements. Results indicated that the factor UTTERANCE TYPE had a significant main effect on the values: $F_{(3,24)}$=12.36, p=.001, η^2=.607. Post hoc comparisons revealed significant differences between statements and wh-questions (p<.001), between wh-questions and declarative questions (p<.001) and between statements and yes-no questions (p<.05). That is, while the wh-questions and the yes-no questions behaved as expected by having significantly higher onsets, the onset values of the declarative questions hardly differed from those of the statements. This pattern was quite robust given that, individually, none of the ten speakers produced significant differences between onsets in the statements and the declarative questions.

As for the factor UTTERANCE LENGTH, onsets turned out to be significantly higher in the short utterances than in the long ones: $F_{(1,8)}$=13.85, p<.01, η^2=.634. In both length conditions, men and women patterned alike, as figure 3.9 shows. The figure also makes it clear that the effect of length held only for the questions, not for the statements, i.e., utterance type and length interacted: $F_{(3,24)}$=5.46, p<.05, η^2=.406. Separate analyses of variance for each question type (with the same set of factors as in the repeated measures analysis, except for UTTERANCE TYPE) showed that the effect of length on onset height was significant in all three question types (WQ p=.024, YQ

p=.007, DQ p=.003).

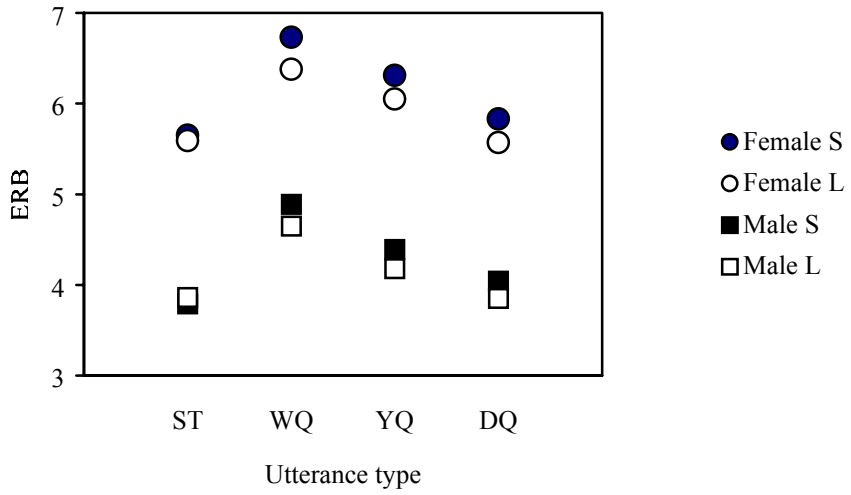


Figure 3.9. Utterance onsets
Mean f0 values of utterance onsets (ERB) in statements and questions for each of the sexes, broken down by UTTERANCE LENGTH (S = short, L = long).

3.5.2.2 Syntagmatic approach

Next, we considered the utterance onsets from the *syntagmatic* perspective. In the sequences of statement-statement and statement-question we established, pairwise, the amount of global reset and local reset by subtracting the onset (or offset) value of the first utterance from the onset value of the second utterance. According to our hypothesis, the amount of reset would be greater between statement and question than between statement and statement (cf. Figures 3.1a,b above). The effect of the factor UTTERANCE TYPE proved significant in both types of reset: Global reset $F_{(3,24)}$=11.85, p<.001, η^2 =.597; local reset $F_{(3,24)}$=10.90, p=.001, η^2 =.577. In local reset, there was also a small effect (just below the .01 level) of UTTERANCE LENGTH: $F_{(1,8)}$=8.31, p=.020, η^2 = .510. That is, the long questions showed less reset than the short ones, which ties in with the finding that the long questions started from a lower f0 (see above) and were less steep (see below) than their short counterparts. Post hoc analyses showed the same pattern for both types of reset, i.e., differences were significant for ST~WQ (p<.001), ST~YQ (p<.05), and WQ~DQ (p≤.005). The absence of significant differences in reset between ST and DQ is in line with the earlier finding that onset heights in these two utterance types were not significantly different. In neither type of reset did the factor SEX OF SPEAKER produce significant effects.

Figure 3.10 displays the results for both types of reset, pooled over SEX OF SPEAKER and UTTERANCE LENGTH. Negative bars reflect a downward step in pitch

between utterance A (which was always a statement) and utterance B (which was either a statement or one of the three question types, cf. the x-axis), pointing to incomplete reset; positive bars reflect an upward step in pitch between the onset of A and the onset of B, that is, reset is overcomplete.

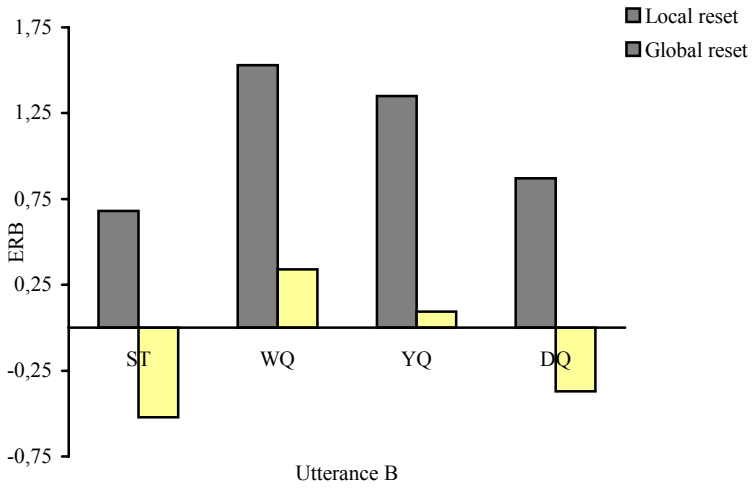


Figure 3.10. Reset
Degrees of local reset and global reset (ERB) in paired sequences of statement-statement vs. statement-question (see text).

Obviously, the two patterns were quite similar in that reset was always greater in the questions than in the statements. Thus, while in a sequence of two statements the onset of the second statement was higher than the offset of the first statement ('local reset'), pitch stepped up higher as soon as the second utterance was a question. When the onset of utterance B was compared with the *on*set of utterance A ('global reset'), pitch stepped up only when utterance B was a wh-question or a yes-no question. As noted above, the finding that global reset did not occur in declarative questions is obviously bound up with the observation that DQ featured relatively low onsets. In chapter 5, we explore whether this acoustic difference between DQ and the other two question types might reflect a functional difference.

Summarising the results on utterance onset height, we can say that the onsets of WQ and YQ were significantly higher than those of ST, whereas DQ patterned with ST (although local and global resets of DQ still exceeded those of ST). It was also found that onsets were systematically higher in the short questions than in the long ones. Although the syntagmatic approach was finer grained in that onset/onset relations and onset/offset relations were determined individually rather than on the basis of overall means, the results of the paradigmatic and syntagmatic approaches were in agreement. The finding that questions following a statement generally displayed higher onset pitch than did statements supported the idea that speakers aim to throw questions into auditory relief, so to speak, presumably to call the listener's attention to the interrogative status of the utterance.

3.5.3 Register level

Register level was taken to be reflected by the minimum f0 value in a given utterance. These minima significantly varied as a function of UTTERANCE TYPE ($F_{(3,24)}$=13, p<.001, η^2=.619). Post hoc tests indicated that the mean register level of each of the three question types differed significantly from that of the statements (ST~WQ: p<.005, ST~YQ: p<.001, ST~DQ: p<.001). Especially in the women, there was a clear dichotomy: the minima in the statements were lower than the minima in the set of questions. The male speakers displayed a more gradual trend, which had the hypothesised order ST<WH<YQ<DQ. Among the three question types, no significant differences were found. In sum, the questions could be said to display higher register levels than the statements. Figure 3.11 illustrates the results.

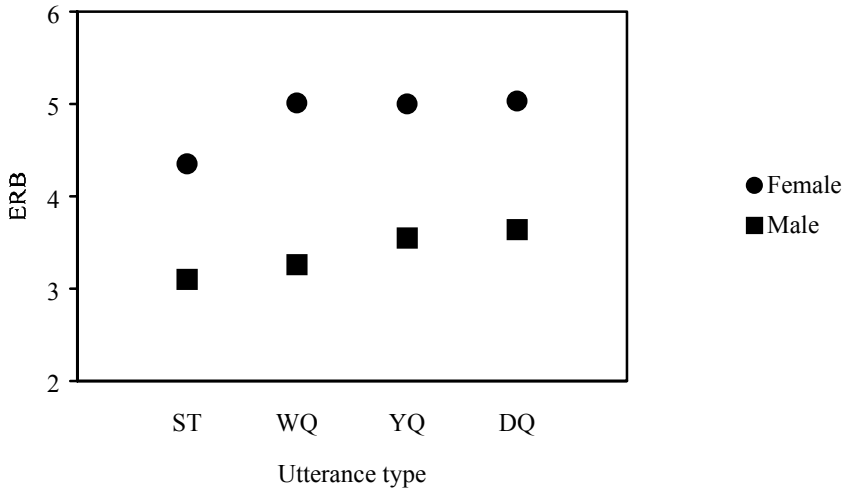


Figure 3.11. Register level (f0 minima)
Means (ERB) of f0 minima broken down by UTTERANCE TYPE and SEX OF SPEAKER.

3.5.4 Global trend

The global trend parameter was investigated with the help of measures considered more objective than declination lines fitted by eye. Thus, on the basis of all-points regression lines that had been (semi-)automatically[18] calculated, lower and upper regression lines were determined. The factor UTTERANCE TYPE had a significant main effect on the trends of the upper and lower regression lines. In the statements and the wh-questions,

[18] As may be recalled, final rises had been excluded from the calculations of the regression lines.

both the upper and lower regression lines were falling; as a matter of fact, in WQ downtrends were considerably steeper than in ST. By contrast, in the yes-no and declarative questions, the upper regression lines displayed clear upward trends; also, in DQ the lower regression line was horizontal rather than downward. Post hoc analyses revealed significant differences between the statements and the declarative questions, between wh-questions and declarative questions, and between yes-no questions and declarative questions. Table 3.3 gives the results of the analyses of variance.

Table 3.3. Global trends
Main effects of the factor UTTERANCE TYPE on upper and lower global trend lines

f0	dF	F	p	η^2	Post-hoc comparisons					
					ST WQ	ST YQ	ST DQ	WQ YQ	WQ DQ	YQ DQ
Upper line	3,24	17.11	.000	.681			.008		.000	.014
Lower line	3,24	13.65	.000	.630			.001		.000	.000

As for UTTERANCE LENGTH, this was found to have a significant main effect on the slopes of the all-points lines only: $F_{(1,8)}=18.47$, $p<.01$, $\eta^2=.698$. That is, in the short versions of ST, WQ and YQ downtrend was steeper than in the long versions; similarly, the upward trend in short DQ was less steep than in long DQ. Separate analyses of variance for the three question types (with the same factors as in the repeated measures analysis, except for the factor UTTERANCE TYPE) showed that the effect of length was significant (at the .01 level) only in the wh-questions ($F_{(1,24)}=27.17$, $p<.001$). However, considering that the all-points lines lacked theoretical status (cf. § 3.3.5) and that length did not significantly affect the upper or lower regression lines, we will not go into this matter any further. In sum, there was no hard and fast evidence that the long questions had shallower slopes than the short ones (cf. § 3.4.1).

To conclude the section on global trends, Figure 3.12 provides an overview of the slopes of upper and lower trendlines, broken down by UTTERANCE TYPE but collapsed over UTTERANCE LENGTH, UTTERANCE POSITION and SEX OF SPEAKER, as these factors did not show significant effects. A negative slope coefficient indicates downtrend, a positive coefficient stands for uptrend.

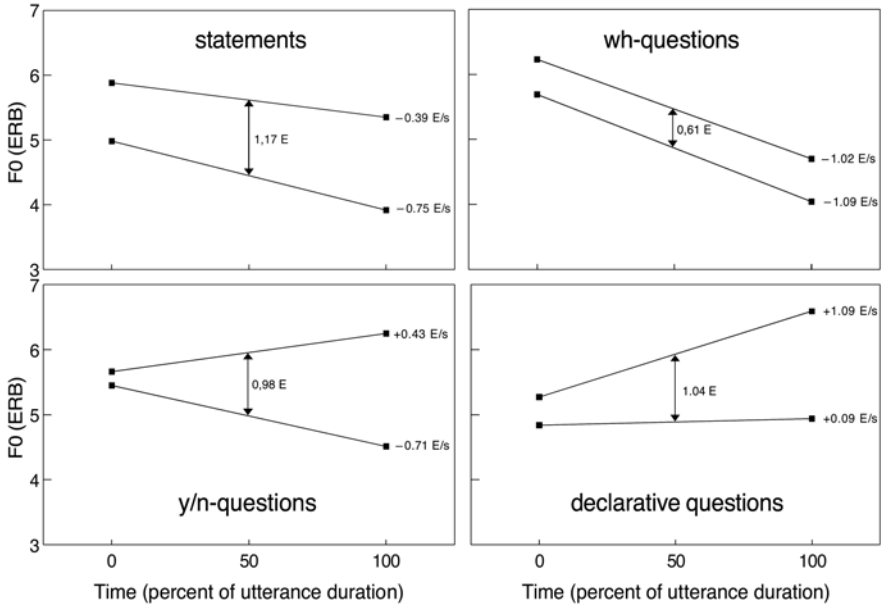


Figure 3.12. Global trends (overview)
Initial and final frequencies (in ERB) of upper and lower regression lines and their respective slope coefficients (in ERB/s) in four utterance types: statements (ST), wh-questions (WQ), yes-no questions (YQ), and declarative questions (DQ); the vertical arrows indicate mean register span at the temporal midpoint.

Summing up the global trends characteristic of each of the four utterance types, Figure 3.12 captures at the same time (relative) register level and (relative) register span. To facilitate the quantitative assessment of the latter, the distance in the ERB domain has been indicated at the temporal midpoint of the utterances (i.e. at the 50% point along the normalised time axis). In WQ, the lower trend line shifts to a very high value, causing register span to be substantially narrowed (see also § 3.14 below).

3.5.5 Utterance Position

The factor UTTERANCE POSITION (that is, did the target utterance occur in first or second position within a pair, or in isolation?) did not show significant main effects or interactions in any of the analyses of variance. This independent variable had been mainly included with a view to monitoring possible differences between the onset heights of the adjacent utterances. Thus, the onset of the *first* utterance, whether statement or question, might well have been relatively high to mark the beginning of a new paragraph. This might then have affected the sizes of possible resets. However, within the sexes and within each utterance type onsets turned out to be fairly constant, as Table 3.4 shows.

Table 3.4. Utterance Position
Means (ERB) of utterance onset values as a function of SEX OF SPEAKER, UTTERANCE TYPE and UTTERANCE POSITION.

Sex of speaker	Utterance type	1st position	2nd position	Isolation
Female	ST	5.66	5.61	5.67
	WQ	6.57	6.52	6.61
	YQ	6.26	6.20	6.02
	DQ	5.70	5.67	5.78
Male	ST	3.86	3.78	3.85
	WQ	4.88	4.72	4.66
	YQ	4.22	4.35	4.28
	DQ	3.90	4.02	3.88

In conclusion, Table 3.5 gives an overview of the significant main effects of each of the four factors (there were no significant second order interactions). Note that the significant effects of SEX OF SPEAKER on the *raw* parameters were due to the intrinsically different pitch levels in male and female speech.

Table 3.5. Main effects: Overview
Significance levels (.01) and η^2 of the four factors SEX OF SPEAKER, UTTERANCE TYPE, UTTERANCE LENGTH and UTTERANCE POSITION. Between brackets: significances at the .05 level.

Factors	Sex		Type		Length		Position	
	p	η^2	p	η^2	p	η^2	p	η^2
Size final rise	(.025)	.485	-	-	-	-	-	-
Onset f0	.000	.939	.001	.607	.006	.634	-	-
Local reset	-	-	.001	.577	(.020)	.510	-	-
Global reset	-	-	.000	.597	-	-	-	-
Minimum f0	.000	.820	.000	.619	-	-	-	-
Maximum f0[19]	.000	.921	.001	.567	-	-	-	-
Upper regression line	-	-	.000	.681	-	-	-	-
Lower regression line	-	-	.000	.630	-	-	-	-
Postnuclear low	.000	.857	.000	.670	-	-	-	-

[19] Maximum f0 only played a role in the calculation of register span needed to compare the sexes, see § 3.6.5 below.

3.5.5 Sex of Speaker

As pointed out earlier, sex-related differences could only be established for the non-raw parameters, given the intrinsic difference between the sexes in pitch level. For the sake of compactness, the results will be presented and discussed in § 3.6.5, following the discussion of the other results.

3.6 Discussion

The questions to be answered in this chapter were the following:

- Are the questions in our production corpus acoustically different from statements in terms of (i) local f0 at the utterances' edges (i.e. utterance onset height and final rise), and (ii) global f0 (i.e. overall trend and register level)?
- Do differences follow the order predicted by the Functional Hypothesis, i.e. ST<WQ<YQ<DQ?

The results, presented in the previous sections, have made it clear that these questions can be answered in the affirmative. In the sections below, we discuss the findings for each of the individual parameters.

3.6.1 Final rises

One of our findings was that, among the intonational characteristics of questions, the final rise had pride of place: it occurred in 520 out of the 600 questions. Excursion sizes did not significantly differ as a function of question type, which was somewhat unexpected in view of the Functional Hypothesis. However, this was fully compensated for by the finding that, in WQ, the final rise was realised lowest in the speaker's range, in DQ highest.

At this point, we should mention two further findings. We observed that, in individual questions, the maximum f0 of the final rise almost always exceeded the question's onset value, even in WQ whose utterance onsets were relatively high. Also, mean maxima of the final rises exceeded mean peak values of the preceding object accents, in the male as well as in the female speakers. These observations suggested a basic strategy in speakers to give their questions an overall upward sweep. Interestingly, Uldall (1962) showed that (American English) utterances were more often classified as questions when the highest pitch of the entire utterance occurred at the very end. Similarly, Studdert-Kennedy & Hadding (1973:304) found that Swedish and American listeners preferred question contours where the end frequency (i.e. of the final rise) was the highest in the utterance. Our production corpus provided clear evidence of this mechanism. We termed the upward step in pitch between question onset and rise offset 'upsweep'. Given that there were three question types and ten speakers and, consequently, 30 (pairs of) mean values, in 29 of these were the maximum frequencies of the final rises higher than those of the questions' onsets, i.e. Δ(e-b) was positive (where 'e' stands for 'end frequency', 'b' for 'begin frequency'). Upsweep was smallest in WQ, largest in DQ, as Figure 3.13

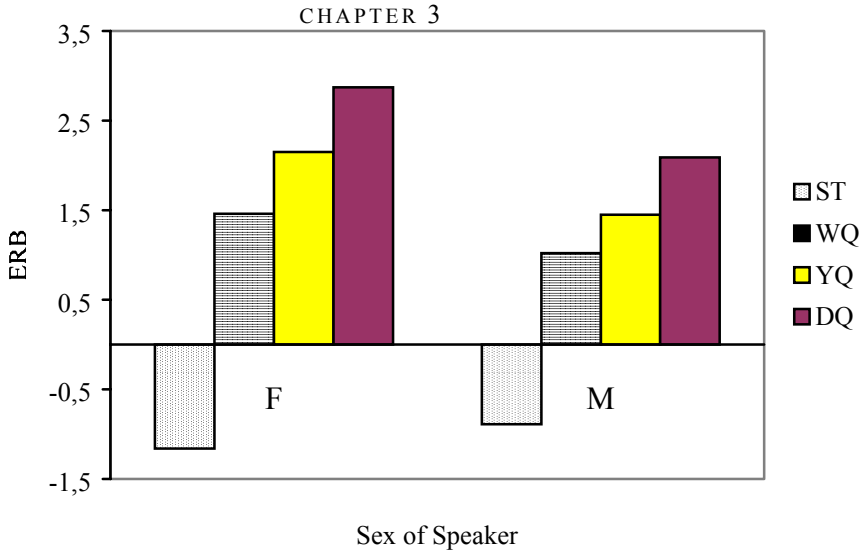


Figure 3.13. Upsweep in questions and downdrop in statements, broken down by SEX OF SPEAKER (F, M).

illustrates. In statements, Δ(e-b) was negative, i.e., there was 'downdrop'[20]. As Table 3.6 below shows, *within* each question type mean Δ(e-b) had a fairly constant size irrespective of the question's length, whereas *across* question types it strongly increased in the expected order WH<YQ<DQ.

Table 3.6. Upsweep in questions
Means (ERB) of Δ (e-b), i.e. the difference between maximum end frequency and utterance onset frequency, broken down by QUESTION TYPE and LENGTH.

Δ (e-b) in:	Short	Long
WQ	.81	.73
YQ	1.47	1.63
DQ	2.19	2.16

In retrospect, the observation that Δ(e–b) in WQ was relatively small would seem to explain why, somewhat unexpectedly, the mean excursion of the final rise was the largest in WQ, the smallest in DQ (although differences between the question types failed to reach significance, cf. § 3.5.1.2). Given that WQ featured the highest utterance

[20] Practical problems forced us to cut out statistical analyses of 'upsweep'/'downdrop'. Breaking down the data into six utterance types rather than four, i.e. ST, WQ+ (with final rise), WQ– (without final rise), YQ+ (with final rise), YQ– (without final rise), and DQ) would have caused many conditions to be absent. As a result, the matrix would have had a large number of empty cells, causing the other data of the speaker to be lost (see § 3.5). In this situation, filling the empty cells with mean values from other cells would have resulted in an unacceptable loss of variance. Moreover, it seems evident that statements will feature downdrop, and that (most) questions will feature upsweep. Hence, it seems trivial to support this with statistics.

onsets together with the lowest rise onsets, the excursion of the final rise had to be large enough to make the offset higher than the question's onset value. Conversely, DQ combined relatively low utterance onsets with relatively high rise onsets. Hence, even a moderate final rise would be sufficient to create upsweep. Yet in this question type, whose questionhood crucially relies on intonation, Δ(e-b) was sometimes well over 2 ERB. As usual, yes-no questions held an intermediate position.

The above observations suggest a strong tendency for questions to meet an upsweep requirement (provided that there is a final rise). To be sure, in WQ a final rise is not indispensable to the interrogative character: this question type may well go without a final rise and, hence, without upsweep. What matters most, however, is that an apparently robust parameter UPSWEEP (cf. Table 3.6) clearly differentiated between the three question types in the order WQ<YQ<DQ, in support of the Functional Hypothesis. And, importantly, *if* WQ features a final rise, it conforms to the upsweep mechanism outlined here.

3.6.2 Utterance onset height.

In two out of the three question types (WQ and YQ), the utterance onsets were significantly higher than in the statements, both from a paradigmatic and a syntagmatic view. These results add to our general understanding of pitch relations among adjacent structures. While no evidence for global reset was found between sentence-internal *clauses* (Collier 1987), or between consecutive *statements* (Sluijter & Terken 1993), global reset did occur in sequences of statements and *questions* (that is, when the latter were WQ or YQ). As pointed out above, the observation that this did not hold for DQ seemed at variance with the Functional Hypothesis. This hypothesis predicted that DQ would typically display the greatest amount of high(er) question pitch, given its lack of lexical and syntactic cues to interrogativity. However, as the previous section showed, the low utterance onset of DQ was compensated for by a large value for UPSWEEP (Δ (e-b)).

In both sexes and in each of the three question types, the short questions had systematically higher onsets than the long ones. This was not true of the statements: here, onsets were found to be quite constant across utterance length and utterance position. Neither of these results was in line with a finding by 't Hart *et al.* (1990:128) that *longer* statements have *higher* initial frequencies. At present, there would seem no ready explanation for this.

3.6.3 Register level

It may be recalled here that the regression model chosen for this study (cf. § 3.3.5) yields trend lines which intrinsically underestimate register level, register width, and steepness of slopes when compared with the declination lines produced by a declination model. This should be kept in mind when the results regarding these parameters are interpreted, or when they are compared with results of experiments based on declination lines. *Within* the present material, however, it is perfectly possible for the global features to be mutually compared.

In agreement with hypothesis (iii), the mean f0 minima were significantly

higher in the questions than in the statements, even in WQ with its relatively steeply downward global trend. This upward shift of register level was taken to confirm the *a priori* impression that questions stand out from neighbouring statements by raised pitch levels. According to Gårding (1982), overall upward (or downward) shifts in register level may express meaningful relations between consecutive utterances, and listeners rely on such cues. For instance, when the two Swedish counterparts of the English statements *She went out* and *She didn't go to bed* imply a contrast, pitch in the latter may display a global upward shift. Our results on utterance onset and register level show that a sequence of a statement and a question may, likewise, display such prosodic discontinuity. Presumably, this serves to highlight the important functional difference between statement and question. As Geluykens (1985) has pointed out, in spontaneous (English) conversation listeners often start to answer a question before the speaker has at all been able to realise a final rise. This led him to speculate that question contours may start from f0 levels higher than the baseline (1987:493), which might then act as an early prosodic cue to interrogativity. Our Dutch production data seem to meet this expectation.

Finally, it may be noted that WQ frequently combined a raised register level with a strongly narrowed register span (Figure 3.12), that is, the distance between f0 minima and maxima was quite small. The f0 curves in Figure 3.14 illustrate this. In the upper panel, a wh-question is preceded by a statement, in the lower panel by a yes-no question (lower panel). The contrast between the two question types YQ and WQ is especially striking. In chapter 4, which deals with the accentuation of the experimental utterances, we explore whether the narrowing of register in this question type may have been due to a specific accent pattern.

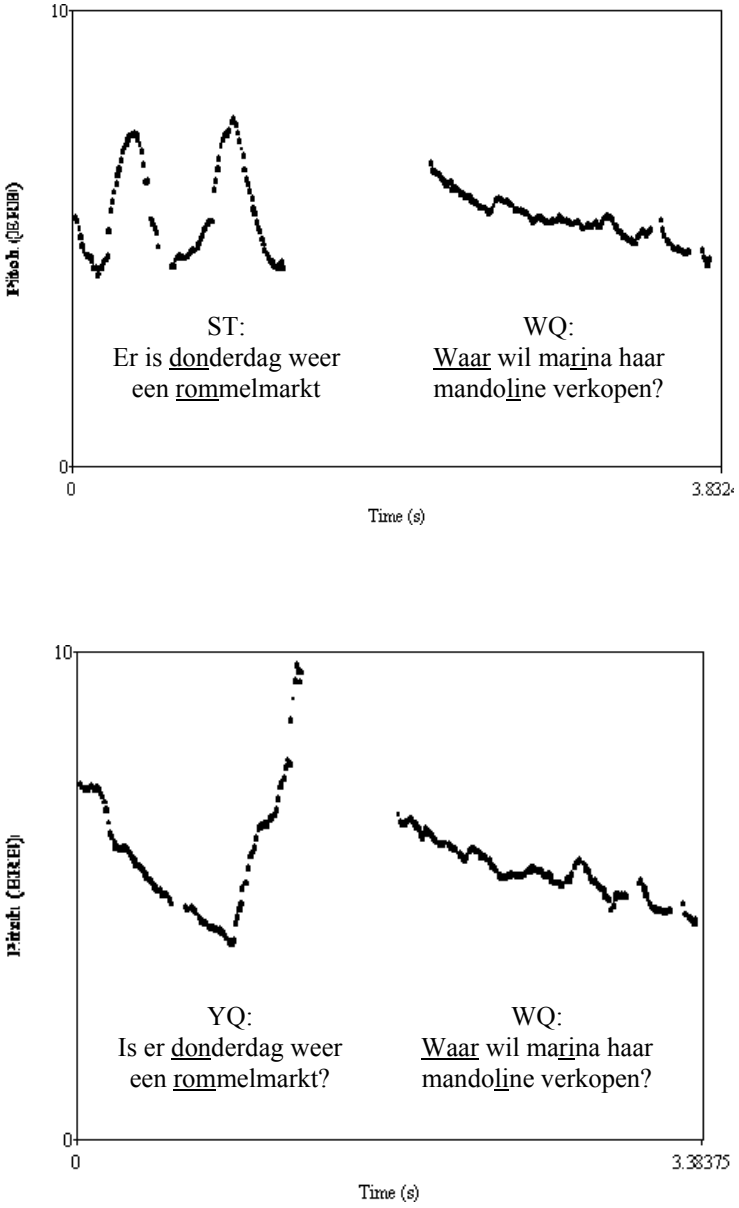


Figure 3.14. Register narrowing in WQ
Sequences of ST + WQ (upper panel) and YQ + WQ (lower panel) realised by speaker RS (female).

3.6.4 Global trend

The reason for comparing and contrasting global f0 trends in statements and questions largely rested on the claim that, in Danish, global f0 slopes vary with utterance function[21]. Thorsen (1978, 1980) found that the slopes of different Danish question types, non-final clauses (representing continuation) and terminal declarative sentences displayed a clearly observable order. At the one extreme, the terminal declarative utterances were the most steeply falling; their slopes were assumed to represent 'unmarked' intonation. At the other extreme were echo questions, whose slopes were horizontal and therefore 'marked'. In between were intermediate slopes of other question types (YQ, WQ, and two others), together with non-terminal clauses (see also chapter 2, § 2.2.2.4.3, and Figure 2.4). Variation among individual speakers did not affect this hierarchy. According to Thorsen, the degree of downtrend depended on the amount of lexical and/or syntactic information: the more non-prosodic information was present, the more falling the intonation contour was, and vice versa (see also Grønnum 1990, 1995, 1998). It seemed worth investigating whether a similar order could be found in Dutch. That is, although the present study uses regression lines rather than the declination lines drawn through the stressed syllables in the Danish study, we nonetheless expected to find the slopes to display the order ST<WQ<YQ<DQ.

However, though this order held for YQ and DQ (the slopes of their lower regression lines ran less steeply downward than those of the statements, and their upper regression lines displayed horizontal and upward trends, respectively), the slopes of WQ were actually more steeply downward than those of the statements. This held for the upper as well as the lower regression lines (cf. Figure 3.12). In other words, the slopes of the three question types failed to cluster together, there was no categorical division between statement and question slopes, as in Danish. Also, there was a tendency for overall trends to vary as a function of utterance length.

Considering this absence of a dichotomy between the statements and the questions, it seemed doubtful for overall slope to be an independently chosen global feature distinguishing utterance types. In this respect, the PL ('postnuclear low') values seemed to make a more appropriate parameter. As may be recalled, in the *statements* the PL was the last reliable f0 value of the entire utterance, whereas in the *questions* the PL represented the onset of the final rise (cf. § 3.4.3 above). In point of fact, it is quite plausible for these PL values to have substantially contributed to the global prosodic distinction between questions and statements. Earlier, we saw that the mean PL values formed an ascending trend following the hypothesised order ST<WH<YQ<DQ. In both sexes, the difference between the statements and the three question types was significant (cf. § 3.5.1.2). Unlike statements, at the end of which f0 is presumably lowered to the speaker's baseline ('final lowering'), questions would seem to aim at keeping pitch relatively high (see also Strik & Boves 1995), possibly because of the production demands of the ensuing final rise. According to Ohala & Ewan (1973, cited in Vaissière 1983:56), it is harder to produce rises than falls. This

[21] Note that there is an *a priori* difference between Dutch and Danish intonation in that the latter has been analysed as having recurrent 'stress groups' of one and the same pattern, rather than a choice from different pitch accents as is the case in, e.g., Dutch and English (e.g. Grønnum 1995). Note, also, that Danish questions lack final rises.

would seem especially true when rises occur at the very end of an utterance and are not followed by a return to the baseline. Our data would seem to support a view that the high PL's in the questions anticipate the final rise. We compared the 280 utterances in the corpus that were realised *without* final rises (i.e. 200 ST, 72 WQ and 8 YQ) with the 421 questions ending *with* a rise. Table 3.7 shows that, in both sexes, mean PL height in the tokens *with* final rises was markedly higher than in the versions ending low (cf. the last column).

Table 3.7. Postnuclear lows (PL)
Mean values (ERB) of PL broken down by Final boundary tone (L% vs. H%) and by Sex of speaker

Final boundary tone	L% (N=280)	H% (N=421)	Δ
Female speakers	4.52	5.49	+ .97
Male speakers	3.18	3.87	+ .69

In order to establish whether this difference was significant, we carried out a pairwise t-test on those speakers who produced both L% and H% in the WQ-subset (i.e. three men, three women). In each of these six pairs, PL was higher before H% than before L%; the effect of boundary tone was significant: $t_{(5)}$=–3.7, p=.007 (one-tailed). Though more research in this area is clearly called for, it is our impression that a projected final rise would seem to maintain a relatively high PL. This, in turn, seems to affect global trend. Thus, as Table 3.8 makes clear, in WQ+ (i.e. *with* final rise, N=125) downtrend was less steep than in WQ– (i.e. when no final rise followed, N=72). The same held for YQ: comparison of the tokens with a low (N=8) and a high (N=165) final boundary tone showed that, before L%, downtrend of the lower regression line was steeper than before H%, even more so than in WQ[22] [23].

Table 3.8.
Mean global trends (ERB/s) of the lower regression line broken down by final boundary tone (L% vs. H%) in WQ and YQ

Final boundary tone	Global trend before L%	N	Global trend before H%	N
WQ	–1.28	70	–.98	125
YQ	–1.70	8	–.77	165

Thus, in retrospect the mean lower regression slopes of YQ in Figure 3.12 appear to have been substantially affected by the eight realisations *without* final rises. Had these been excluded, the slopes would have been less steeply downward. These observations suggest that mean (lower) slope derived at least partially from the PL value, making

[22] Note that occurrences of high H% (H* H%) were excluded here as they would have unduly raised the global trendlines in the H% group. Of the eight occurrences of L% in YQ, four were produced by men, four by women.
[23] It is hard to assess whether the differences in Table 3.8 are significant. A t-test for independent groups yields significant results, but violates the assumption of independent measurements. At the same time, in YQ there were only eight renderings with L%. Because of this small number, the material was unsuitable for a full-blown repeated measures procedure.

slope an indirect rather than a direct cue[24]. In point of fact, the PL values seem to have functioned as hinges. Constituting the *onsets* of the final rises in the order WQ<YQ<DQ, they enabled the *offsets* of these final rises (whose excursion sizes were *not* significantly affected by question type) to exceed the questions' starting frequencies, even if the latter stepped up in the *reverse* order DQ<YQ<WQ. This regularity in the behaviour of the (local) PL's was not revealed by the (global) regression lines (which, anyway, included only the starting points of the final rises, not their end points). The observation that, in our production data, the PL values seemed to constitute a major factor in the questions' acoustics may be related to the perception experiments by Hadding-Koch & Studdert-Kennedy (1964:180) and Studdert-Kennedy & Hadding (1973), discussed in chapter 2 (§ 2.2.2.3). Among other things, the researchers found that, in (Swedish) utterances in which peak f0, f0 of PL[25] and final f0 had been varied, an increase in the f0 of the PL was accompanied by an increase in the number of questions heard.

3.6.5 Sex of Speaker: Results and discussion

Earlier, § 3.3.6 gave an outline of empirically established differences between men and women as regards their speech and their communicative behaviour. On the basis of these, we expected women to feel more disposed than men to ask questions and to make a more pronounced use of prosodic means for signalling interrogativity, specifically in the area of pitch range. Though the literature has often associated female speech with a wider register span[26], findings in this area have not yielded compelling evidence for this. However, these findings were based on *declarative* speech. We expected that, in *interrogative* speech, sex-related differences in terms of register span were more likely to show up, locally as well as globally. Hence we predicted that women would realise more and larger-sized (local) final rises as well as greater (local) resets, and that the directions of their global trend lines would be more upward. As for the newly proposed parameter UPSWEEP, we also want to know whether the women had larger values on Δ(e–b) than the men. In the next two sections, our findings are presented and discussed.

Number of final rises

The wh-question was the only question type where speakers seemed to have a certain freedom to decide whether or not they would produce a final rise. If Bolinger (1978b:502) was right when he posited: "the 'more questioning' a question is, [……], the greater the tendency is toward terminal rise", one might expect women to feel more inclined to put in a final rise, e.g. because they want to emphasise the interrogative status of the utterance even though this question type would not seem to strictly require it. According to Fónagy (1969) cited in Bolinger (1989:23), Hungarian wh-questions "with a final rise were more frequent in women's

[24] When global trend is a derived acoustic property rather than an independently chosen feature of question intonation, this is not to say that it cannot serve as a *perceptual* cue to the distinction between questions and statements (see the experiment reported in Van Heuven & Haan 2001, 2002 to appear).
[25] In Hadding-Koch & Studdert-Kennedy's study, PL was termed 'turning point'.
[26] That is, on a scale that allows male and female speech to be directly compared, such as ERB.

pronunciation", and "the final rise gives the sentence a feminine nuance, the falling melody is felt to be more masculine".

In the overall number of 200 WQ, the female speakers produced 78 final rises against 50 in the male speakers, a difference that may be taken to support our expectation (see Figure 3.4b above). It will be obvious, however, that more experiments have to be carried out to determine whether presence or absence of a final rise in this question type is motivated by a speaker's sex or, rather, by more general pragmatic factors. As is well known, a final rise may render a wh-question more friendly, or more polite (cf. Hirst & Di Cristo 1998:26), in men as well as in women. At present we can only say that, *if* final rises in WQ are fully optional, women seem more inclined to produce them.

Excursion sizes of final rises

As Figures 3.7 and 3.8 above illustrated, the final rises in the female questions turned out to be larger than those in the male questions. However, the difference was significant at the .05 level rather than at the pre-set .01 level. For this reason we have to regard the phenomenon as a strong tendency rather than as an established fact.

Utterance onset f0 and reset

Given that the parameter ONSET HEIGHT was expressed in raw values, it was not possible to establish whether the higher onsets of the female speakers were solely due to the biological difference or whether they were, to some extent, gender-controlled. As regards the sizes of local and global reset, the factor SEX OF SPEAKER did not have significant effect on either of them.

Register level

The parameter REGISTER LEVEL was also expressed in raw values, precluding comparison between the male and female speakers.

Global downtrend

As Table 3.5 above showed, there were no significant effects of the factor SEX OF SPEAKER on this parameter. That is, in the female speakers slopes were not systematically less downward, or more upward, than in the male speakers.

Register span (or: 'pitch range')

'Pitch range' was defined as the difference between the mean maximum f0 and the mean minimum f0. Although a Manova (repeated measures design) indicated that the effect of SEX OF SPEAKER on register span fell just short of significance[27], Table 3.9 shows that the mean differences were yet appreciable.

[27] Results of the Manova were $F_{(8,1)}=5.31$, $p=.05$, $\eta^2=.399$, i.e., significance fell below the .01 level.

Table 3.9. Mean register span (ERB) of five female and five male speakers broken down by Utterance type (final rises were excluded from the calculation). Between brackets: standard deviations.

	Female speakers	Male speakers	Δ Female/Male
ST	2.95 (1.03)	2.03 (.48)	+ .92
WQ	2.36 (.53)	2.20 (.64)	+ .16
YQ	2.61 (.69)	1.98 (.54)	+ .63
DQ	2.73 (.85)	1.96 (.54)	+ .77
Mean	2.66	2.04	+ .62

Somewhat surprisingly, the difference between the sexes was greatest in the statements, where it amounted to almost a full ERB. As for WQ, it turns out that the considerable narrowing of register found in this question type was in fact due to the female speakers; in the men, register in WQ was relatively wide. It should be pointed out, however, that individual speakers showed quite some variation here; probably, the phenomenon of narrowed register in WQ is, to some extent, speaker-specific.

Upsweep and downdrop

Figure 3.17 displays the results for the new parameter Δ(e–b) broken down by SEX OF SPEAKER. Negative bars reflect that utterance-final f0 was lower than utterance-initial f0 ('downdrop'); positive bars indicate that utterance-final f0 was higher than utterance-initial f0 ('upsweep'). The labels 'WQ–' and 'YQ–' represent the realisations without final rises, 'WQ+' and 'YQ+' the realisations with final rises. In the female speakers, all utterance (sub)types typically show surplus values, that is, the negative values are lower whereas the positive values are higher than in the male speakers.

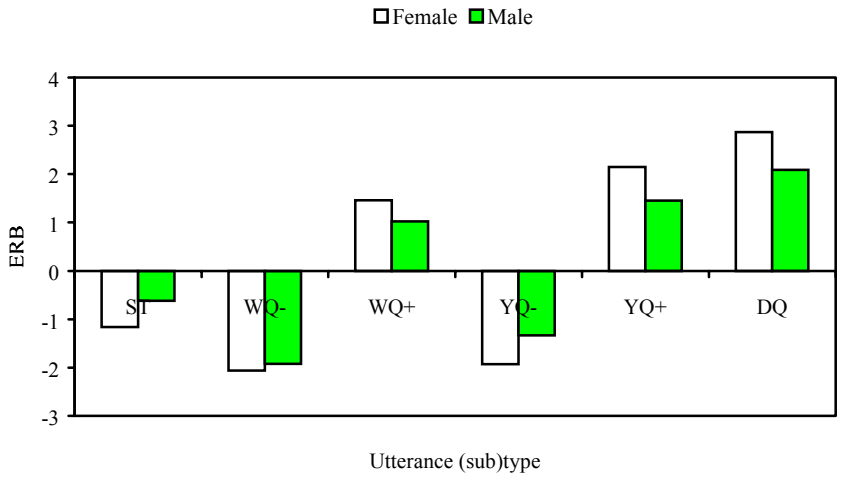


Figure 3.17. Upsweep/downdrop, broken down by SEX OF SPEAKER.
Δ(e–b) in statements, in WQ without final rises, in WQ with final rises, in YQ and in DQ.

Assuming that Δ(e–b) is one of the prosodic devices available to speakers, the women clearly put it to a more pronounced use. In their speech, the various utterance types were prosodically more spaced-out, the questions as well as the statements. This resulted in greater contrasts between the various discoursal and attitudinal categories as expressed by ST, WQ+ vs. WQ– , YQ+ vs. YQ–, and DQ. We interpret this as evidence for the greater expressivity that is generally attributed to female speech (cf. § 3.3.6.1 above) and which, presumably, enhances women's communicative effectiveness.

Discussion and conclusions

Considering that the effect of SEX OF SPEAKER was significant at the level of .05 rather than .01, we conclude that there is a *tendency* for female speech to have a globally wider register span. Our expectation that this would show up in, particularly, questions was not confirmed, though. Likewise, on the new parameter Δ(e–b) the women's values were found to exceed those of the men in the statements as well as in the questions. Assuming that Δ(e–b) can be viewed as a global reflection of register span, too, this meant that female span was, again, wider, across the four utterance types.

Furthermore, not only did the women make more (local) final rises in their questions, these rises had also larger excursion sizes (although, again, the significance of .025 fell below the level of .01). Thus, in aggregate the results did not confirm our expectations that wider register span in female speech would show up in local rather than in global measures, and in questions rather than in statements. In fact, it showed up locally as well as globally, and in statements as well as in questions, however without meeting the significance level of .01.

Our finding that, on average, the women were more disposed than the men to expand their pitch range would seem to detract from Henton's (1989) claim that the alleged 'swoopiness' of female speech is a stereotype rather than an empirically established phonetic fact. Henton's study, where 'swoopiness' was taken to be acoustically reflected by 'a greater overall pitch range', did not yield clear evidence of such a difference between the sexes. The present study shows that female speech featured both local and global differences in pitch range. These differences were not limited to the questions, as expected, but were also observed in the statements. Increasing the contrasts between the six utterance (sub)types under investigation, they lent the female speech greater prosodic expressiveness. As this property is likely to put women at a communicative advantage, there seems no reason for viewing it in anything else than in a positive light.

3.7 General summary

The present chapter reported a production experiment involving three Dutch question types whose acoustic characteristics were compared with those of corresponding statements. As a first step, we concentrated on global pitch phenomena (REGISTER LEVEL and GLOBAL TREND), as well as on local pitch phenomena associated with the utterances' edges (ONSET HEIGHT and FINAL RISE). Given crosslinguistic reports that questions are typically characterised by high(er) pitch somewhere in the utterance, we

hypothesised that our experimental questions would display higher utterance onsets, raised register levels, less downtrend, and utterance-final rises. In addition, a Functional Hypothesis predicted that these acoustic properties would be more pronounced in question types lacking one or both of the non-prosodic markers of interrogativity (i.e. subject-verb inversion and/or a question word). Hence, we expected the three question types to display the order WQ<YQ<DQ. Moreover, on the basis of claims about differences between male and female speech and communicative behaviour, the female renderings of the experimental utterances were hypothesised to feature larger pitch intervals than the male renderings, particularly in the questions. By and large, our results supported these hypotheses, i.e.,:

- Questions showed a predominance of final rises (86.6% across the three question types);
- In the questions, utterance onsets, register levels and global trends featured higher pitch than in the corresponding statements;
- Pitch levels generally followed the order WQ<YQ<DQ, providing ample support for the Functional Hypothesis;
- Female speech displayed an overall tendency towards a wider register span, globally as well as locally, and not only in the questions but also in the statements.

This is not to say, however, that all results allowed equally smooth interpretations. *Prima facie*, two observations appeared to go against the Functional Hypothesis. First, the overall trends of the statements and the three question types failed to display the expected order ST<WQ<YQ<DQ. Besides, trends were also observed to vary as a function of onset f0, utterance length, and presence vs. absence of a final rise. These observations made it highly unlikely for global trend to represent an independent speaker choice, varying with discourse function. Rather, they suggested that global trend is a secondary, derived feature. Second, mean utterance onsets of DQ were not significantly higher than the mean onsets of ST, WQ or YQ.

However, these apparently deviant results were compensated for by additional findings. First, PL value (the 'postnuclear low') seemed a more appropriate parameter along which the four utterance types could be distinguished, than overall trend. PL values stepped up in the order WQ<YQ<DQ (as predicted by the Functional Hypothesis), more or less in tandem with the maximum values of the final rises ('rise offsets'). This mechanism, which could not be captured by the automatic trend lines, caused the final rises to be realised higher in the speaker's range as a function of question type. Also, the results suggested a certain interdependence between a question's initial f0 and final f0 (when a final rise was present). Thus, across individual speakers and question types, f0 at the end of the rise turned out to be systematically higher than f0 at the beginning of the question, that is, Δ(e–b) was positive. Apparently, speakers aimed at placing the highest pitch of the entire utterance in the final portion, thereby giving the questions a distinct upward sweep ('upsweep'). In this process, the upstepping PL values seemed to act as hinges. Values on the new parameter Δ(e–b) strongly increased in the order WQ<YQ<DQ, again fully in line with the Functional Hypothesis. Interestingly, on this parameter the women produced surplus values across the board, that is, their

positive values were higher than those of the men, their negative values lower, in the statements as well as the questions. This greater prosodic differentiation between the various utterance types was taken to reflect greater expressiveness, in line with claims to this effect in the literature. Presumably, this greater degree of expressivity enhances women's communicative effectiveness.

The findings so far indicate that speakers have a variety of devices to raise pitch when creating question contours:

- Utterance onset and minimum f0 can be raised to create prosodic discontinuity between a question and preceding statements; this, presumably, serves to highlight the functional difference between the two;
- A final rise can be inserted; projection of this rise may cause PL values to be raised which, in turn, may cause overall trends in questions to become less downward;
- The maximum f0 of the final rise can be made to overstep the f0 of the utterance onset ('upsweep'), a mechanism that further enhances the overall impression of 'upwardness' in questions.

Although the results were on the whole straightforward, it must not be ignored that variation among individual speakers was sometimes considerable. Apparently, there is a fairly broad latitude in what may be considered a canonically intonated question. The observed variation also confirmed that, for the description of intonation, it might be hazardous to rely on just one or two speakers of a language or variety.

The next part of the investigation concentrates on accentuation. Our main object is to test hypothesis (iv), which predicted that questions feature raised nuclear accents. Also, given that the data provided no evidence that global trend is an independent property of interrogativity, we mean to explore whether and to what extent the global regression slopes were determined by position and realisation of local accents (cf. Bruce 1982b:113; Ladd 1983:47; Liberman & Pierrehumbert 1984:224; Eady & Cooper 1986:412). Therefore, in chapter 4 our investigation of the corpus is extended to include these aspects.

Chapter 4

Production (2): Questions and accentuation

4.1 Introduction

In the previous chapter, a substantial part of the acoustic characteristics of the three selected question types was investigated. Our first approach included the global parameters REGISTER LEVEL and GLOBAL TREND, as well as the edge phenomena ONSET HEIGHT and FINAL RISE. Results were fairly straightforward: on the whole, they indicated that the chosen f0 parameters systematically varied as a function of utterance type. The present chapter looks at the *internal* structures of these global stretches. The primary objective is to test hypothesis (iv), which predicted that the object accents in the set of questions are raised (relative to those in the corresponding statements). This, obviously, requires a detailed investigation of the distribution and shapes of the pitch accents, their excursion sizes and the heights of their peaks. First, for the *qualitative* assessment of the material a transcription is needed to identify possible phonological differences between the statements and the three question types, as well as between the question types themselves. Crucially, this transcription also opens up the possibility to carry out analyses across phonologically homogeneous subsets. For instance, when we calculate the mean peak height of pitch accents, we are able to keep the two different accent types H*(L) and L*(H) strictly apart. Next, for a *quantitative* assessment we need f0 measurements of all relevant accentual pivot points to carry out statistical analyses and to establish whether differences between the four utterance types are significant.

Secondarily, the finer-grained accent data can help us to establish whether there is a correspondence between the observed global regression slopes (chapter 3) and the local f0 events of accentuation. As the results in chapter 3 made it unlikely for the slope differences between the utterance types to have rested on independent speakers' choices, these slopes may have reflected differences between the four utterance types as regards position and realisation of accents. Again, the transcription results offer the opportunity to exclude (most of the) phonological variation and concentrate on phonetic differences between the respective utterance types.

In sum, the present chapter addresses the following two research questions:

- Are the nuclear accents in our question material raised relative to the corresponding accents in the statement material (hypothesis iv)?

- Is it possible to explain the differences in global slope observed in chapter 3 on the basis of the accent patterns typical of each of the four utterance types?

The chapter is organised as follows. To start with, § 4.2 reports on the results of the transcription of the corpus material. After some preliminary remarks in § 4.2.2.1, §§ 4.2.2.2-3 present and discuss the paradigmatic results, that is, the distribution of individual transcription labels across boundary tones and pitch accents. Next, §§ 4.2.2.4-5 deal with the syntagmatic results, that is, they identify accent patterns of entire utterances. This provides us with various canonical patterns of the four utterance types, ranging from the basic, unmarked ones to the more idiosyncratic versions. Proceeding to the quantitative aspects, § 4.3.1 motivates the selection of parameters and § 4.3.2 deals with the statistical analyses. In § 4.4, the results of these analyses are summarised; in § 4.5, they are discussed against the background of the two research questions underlying this chapter. Finally, § 4.6 provides a general summary.

4.2 Qualitative aspects

4.2.1 Transcription of the material

The corpus of read speech was transcribed by the author using the recently developed ToDI system (*Transcription of Dutch Intonation*, Gussenhoven 2002b, to appear). Table 4.1 lists the phonological contrasts proposed in ToDI; 'H' and 'L' point to a high and a low tone, respectively, '%' refers to tones at boundaries, and '*' indicates that the tone associates with the accented syllable. There are three initial boundary tones and two (optional) final boundary tones[1]. The singleton basic pitch accents H* and L* represent level stretches of pitch, i.e., they 'spread', or form pitch plateaus. The pitch accents H* and H*L may be modified, that is, '!' means that the accent is downstepped and '+' reflects a steep (prefinal) fall rather than a gradual one. Finally, the prefix L* indicates that the realisation of the starred tone is 'delayed', that is, the star associates with a post-accentual syllable rather than with the accentual syllable.

[1] As high and low final boundary tones are optional there is a third label, the single '%' (i.e. without H or L), which refers to the *absence* of a separate tonal specification for the boundary. It indicates that the speaker ends on the tone specified by the last tonal target, or that the speaker ends midway a fall or a rise ('half-completion').

Table 4.1. ToDI: transcription labels reflecting phonological contrasts.

INITIAL BOUNDARY TONES	BASIC PITCH ACCENTS	MODIFIED PITCH ACCENTS		FINAL BOUNDARY TONES (OPTIONAL)
%L low	H* high-level	!H* downstepped		L% low
%H high	L* low-level			H% high
%HL falling	H*L falling	H*+L steeply falling (prenuclear)	!H*L downstepped	
		L*HL delayed	L*!HL downstepped + delayed	
	L*H rising			
	H*!H 'chanted call'			

To this set, we added the diacritic '^' to identify cases in which the second accent was perceived to be upstepped relative to the first. This was necessary, considering that hypothesis (iv) predicted that the nuclear accent (i.e. the accent on the object) would be raised. It was also possible for this diacritic to indicate that the object accent peak stepped up relative to the H of a preceding L*H pitch accent or, in the absence of a pitch accent on the subject, relative to a high or high-falling initial boundary tone (analogously to ToDI's downstep), or relative to a high or high-falling accent on the (initial) wh-word. Furhermore, we extended the use of the diacritic '+' to all falls, whether prefinal or final, to record possible differences between statements and questions on this count. It must be emphasised that the added labels were not, *a priori*, expected to reflect phonological contrasts.

The statements (ST), yes-no questions (YQ) and declarative questions (DQ) were assumed to have two accent slots, one for the subject (i.e. *Renée* and *Marina*), and one for the object (i.e. *vlees* and *mandoline*). The wh-questions (WQ) were expected to have an additional accent slot for the question word (i.e. *Wat?* and *Waar?*). Altogether, the material had 1800 potential pitch accents.

Labels were primarily assigned on the basis of auditory information; pitch traces played a supplementary role. The label assignment was spot-checked by two experts, CG and VH, who were presented with 50 utterances evenly spread out over the four different utterance types. In one half of the cases (25), the author expected labelling to be a matter of routine. For the other 25, she had selected utterances whose transcription was not immediately obvious. In these 'problem contours', it was usually just one or two labels that gave cause for doubt, rather than the utterance as a whole. Across the 50 test utterances, both transcribers assigned 212

labels, which were subsequently compared with those of the author. Table 4.2 gives the results for the supposedly unproblematic cases ('routine labels') and the doubtful cases ('problem labels'). 'A+A+A' indicates that all three transcribers agreed; 'A+A+B' points to disagreement in one transcriber, and 'A+B+C' reflects that each of the three transcribers assigned a different label.

Table 4.2. Transcription comparison
Percentages of (dis)agreement among three transcribers of 50 utterances containing 212 (potential) ToDI labels.

(DIS)AGREEMENT	'ROUTINE LABELS'		'PROBLEM LABELS'		TOTAL	
A+A+A	81	78%	60	55.5%	141	66%
A+A+B	23	22%	40	37%	63	30%
A+B+C	0	0%	8	7.5%	8	4%
TOTAL	104	100%	108	100%	212	100%

Obviously, in the majority of cases the three transcribers agreed (A+A+A: 66%). As expected, full agreement mostly involved the 'routine labels'. As regards the category (A+A+B), the cases which the author had previously considered 'routine' proved indeed unproblematic given that 22 out of her 23 labels agreed with the labels of one of the experts. This suggested that her labelling of the non-spotchecked routine cases could be viewed with a fair amount of confidence. As for the category (A+A+B) in the 'problem labels', here each of the three transcribers was the disagreeing party in roughly one-third of the cases. Finally, the eight most doubtful cases (A+B+C) were resolved by the author after careful consideration of the three different transcriptions. In aggregate, most disagreements concerned (i) the presence vs. absence of an accent on the subject, (ii) the type of pitch accent, and (iii) the type of initial boundary tone. On the labels for final boundary tones, the transcribers always fully agreed. In quite a few cases, accenthood gave the impression of being signalled by lengthening only. However, considering that the present study was concerned with the melodic aspects of question intonation rather than with other prosodic phenomena such as duration or amplitude, only pitch events were labelled.

4.2.2 Results

4.2.2.1 Preliminary remarks

As in the previous chapter, results are split up in a syntagmatic and a paradigmatic part. First, §§ 4.2.1.1-5 present the paradigmatic results, that is, each of the categories 'initial boundary tone', 'accent on wh-word', 'accent on subject', 'accent on object', and 'final boundary tone' is dealt with separately. Next, §§ 4.2.1.6-9 look at the results from a syntagmatic angle by indicating which sequences of labels were characteristic of each of the utterance types. Prior to this, however, we have to address the issue of the non-neutral realisations. At times, some of the speakers alternated neutral renderings of the utterances with readings that had somehow an emphatic ring, presumably in a commendable attempt to avoid monotony (it may be recalled that they had been

instructed to read in a lively manner, pretending they were actors in a radio play). Although, at this point, a discussion of the relation between accent pattern and focus in questions is somewhat premature[2], we will nonetheless touch briefly on the subject in order to explain why some of the apparently emphatic contours were discarded whereas others were not.

The two core utterances in our corpus, each of which featured a subject followed by an object, commonly had accents on both subject and object. Sometimes, however, the subject was deaccented. Crucially, deaccentuation does not alter a proposition. Instead, it reflects the discourse status of the constituent, such that the speaker believes this constituent to be already present in the addressee's consciousness. Accordingly, we interpreted realisations in which the subject was left unaccented as 'neutral'. However, in a number of utterances speakers accented *only* the subject. This accentuation clearly affected the interpretation. It caused the subject to be implicitly contrasted with other potential subjects, such that one could felicitously insert the phrase "rather than…" (cf. Chafe 1976). As (4.1) makes clear, without a proper context such contrastive readings are pragmatically incongruous (target utterances are underscored, capitalised words carry the accent).

(4.1) The cat needs feeding. *Does RENEE have some meat left?*

The accent pattern of (4.1) would fit contexts such as (4.2) below, where *Renée* is explicitly contrasted with other people who might be able to supply meat for the cat.

(4.2) We'll need some meat to feed the cat. I asked Emma, but her fridge was empty. Tom only had fish, no meat whatsoever. *Does RENEE have some meat left?*

Considering that realisations such as (4.1) unambiguously forced a non-neutral interpretation (cf. Chafe 1976), it was decided to leave them outside the investigation, along with one token in which the only accent was carried by the verb.

Alongside these obvious examples of contrastive readings, there were two further types of accentuation that sounded not entirely neutral. First, in some of the wh-questions the accents on the wh-words *Waar?* or *Wat?* tended to sound relatively emphatic. However, this was a matter of degree, rather than a categorical shift in interpretation, as was the case with the subject accent above. Hence, there seemed no principled reason for excluding these tokens. Indeed, of some languages it is reported that the question word is standardly realised with emphasis. Second, in one token it was the accent on the *object* that was relatively salient. Now, in a neutral reading of a subject-verb-object utterance it is common for the object accent (i.e. the nuclear accent) to be perceived as slightly more salient than the subject accent. In fact, this may be regarded as the default situation. However, the salience of the object may be increased to signal contrastivity, conveying that the object is being contrasted with other potential objects. Consequently, when a nuclear accent sounds more prominent than a prenuclear accent there is inherent ambiguity between a neutral reading and a contrastive reading

[2] This issue is addressed in chapter 6.

(e.g. Chafe 1976:36; Taglicht 1982:213; Eady and Cooper 1986:402; Ladd 1996:163; Gussenhoven 1999a:44). In the case of our experimental statement *Marina wants to sell her mandolin*, this would mean that the object *mandolin* is implicitly contrasted with other relevant objects such as, say, a flute or a guitar. Although, in theory, one might expect that this ambiguity is removed when the prominence of the nuclear accent increases beyond some critical value, it has proved difficult to establish how much prominence has to be added to bring about a categorical shift (cf. Ladd & Morton 1997). As there is no principled way to draw the line here, we decided not to exclude the (single) utterance suggesting an emphatic realisation of the object. Table 4.3 shows the number of contours excluded from further consideration.

Table 4.3. Distribution of non-neutral realisations of utterances.

Utterance type	Contrastive accent on:		Neutral	Total
	Subject	Other (i.e.verb)		
Statements	0	0	200	200
Wh-questions	13	1	186	200
Yes-no questions	1	0	199	200
Declarative questions	2	0	198	200
Total	16	1	783	800

4.2.2.2 Paradigmatic results

4.2.2.2.1 Initial boundary tones (%T)

At times, deciding on the right labels for the boundary tones in the wh-questions proved somewhat troublesome: the circumstance that the utterance-initial, accented wh-words (*Wat?/Waar?*) were monosyllabic made it hard to distinguish between boundary tone and accent tone. Nonetheless, boundary tones in WQ were mostly labelled %L as, in a majority of cases, an initial rise could be distinguished, even though this was usually very short. This assignment of %L was in line with intuitions that preposing a monosyllabic word such as *En* ('And') or *Maar* ('But') before the wh-word would have resulted in a low rather than a high boundary tone. Thus, the earlier finding that, in general, question onsets were phonetically higher than statement onsets (cf. chapter 3, § 3.5.2) was not reflected in a greater incidence of the phonological category %H. In point of fact, the observed higher onsets may well have been due to the overall upward shift of register level typical of the questions.

Table 4.4 gives the distribution of the three initial boundary tones across the four utterance types. Note that the occurrence of %HL was almost exclusively due to one (female) speaker. Otherwise, the distribution of the initial boundary tones did not show a sex-based bias.

Table 4.4. Initial boundary tones (%T): Distribution

%T in:	%L	%H	%HL	Total[3]
Statements	191 (96%)	0 (0%)	9 (4%)	200
Wh-questions	170 (91%)	16 (9%)	0 (0%)	186
Yes-no questions	160 (80%)	27 (13%)	12 (6%)	199
Declarative questions	190 (96%)	3 (1%)	5 (2%)	198
Total	711 (91%)	46 (6%)	26 (3%)	783

4.2.2.2.2 Accents on wh-words

In 185 out of the 186 neutral wh-questions, the wh-word was perceived as accented. Accents were predominantly high-falling (132); if anything like an accentual rise leading up to H*(L) was present it was usually fairly short, particularly in comparison with the often extensive fall. Half of the occurrences of high-falling accents were classified as H*+L, that is, as steeply falling; in the others, the fall was more gradual. In 53 of the wh-questions, accentual pitch remained high/level, forming a plateau between the initial and subsequent accent. Table 4.5 shows the distribution.

Table 4.5. Wh-words
Distribution of accent types.

	High-falling		High-level	No accent	Total
	H*L	H*+L	H*		
WQ	64	68	53	1	186

The distribution of labels over the wh-words did not show systematic differentiation as a function of SEX OF SPEAKER.

4.2.2.2.3 Accents on subjects

In the 200 neutral statements, the subject accent was realised in 199 cases. This was in strong contrast with the wh-questions, where the subject was *de*accented in 163 out of the 186 neutral realisations. In the 199 neutral yes-no questions, deaccentuation occurred in 32 tokens, in the 198 neutral declarative questions, in 6 tokens. Table 4.6 gives the distribution of accent types across the four utterance types.

[3] Recall that, out of the original total of 800 utterances, 17 tokens were discarded because of contrastively realised accents.

Table 4.6. Subject accents
Distribution of accent types broken down by utterance type.

	High-falling		High-level	Low	No accent	Total
	H*L	H*+L	H*	L*(H)		
ST	78	71	41	9	1	200
WQ	14	0	9	0	163	186
YQ	105	15	47	0	32	199
DQ	133	22	36	1	6	198
Total	438		133	10	202	783

When there *was* an accent, the high-falling and high-level types obviously predominated. Of the mere ten occurrences of the low accent, nine were realised by the same female speaker. The label L*(H) was used only when the accented syllable was auditorily perceived as fully low (cf. ToDI). Alignment of the accent peak with the segmental material was not taken into account, in contrast to Grabe (1998:58) where the distinction between H*L and L*H was interpreted in terms of the position of the starred tone relative to the accented syllable[4]. Thus, in Grabe's pitch accent H*+L "the peak of the rise is reached within the stressed syllable, in L*+H the peak is reached within the following syllable" (p.83). This would seem to suggest that, in L*+H, part of the rise to H occurs within the accentual syllable which, as a consequence, may not be fully low. Also, there are indications that peak alignment may vary with discoursal function. Thus, questions have been found to have later peaks than statements, and utterances with later peaks have elicited more question responses (e.g. Gósy & Terken 1994, Makarova 1999a, Remijsen & Van Heuven 1999, Gussenhoven & Chen 2000, Van Heuven & Haan 2001)[5]. Hence, in our transcription alignment did not enter into the distinction between high and low accents.

4.2.2.2.4 Accents on objects

In majority, objects were accented; the 69 non-occurrences of object accent were found exclusively in WQ. Together with the earlier observation that, in WQ, the subject was deaccented even more frequently (in 163 cases), this suggested that accentual prominence in WQ was largely located in the utterance-initial wh-word.

Similarly to the subject accents, object accents were predominantly high-falling (or high-level). In the objects, the high-level accents were always part of the

[4] See also § 4.2.2.4.5 below.
[5] Alignment of H*L peaks has also been shown to vary as a function of, e.g., segmental composition of the accented syllable (Volskaya 1999) and the discourse property [+topic] vs. [-topic] (Wichmann, House and Rietveld 2000; the latter also presents an overview of research results concerning variation in alignment).

'high H%' configuration H* H% (cf. chapter 3, § 3.5, and Figure 4.1b below). Low accents were, again, relatively rare. Table 4.7 gives the full distribution.

Table 4.7. Object accents
Distribution of accent types broken down by utterance type.

	High-falling		High-level	Low	No accent	Total
	H*L	H*+L	H*	L*(H)		
ST	64	136	0	0	0	200
WQ	26	63	3	25	69	186
YQ	7	138	28	26	0	199
DQ	16	107	63	12	0	198
Total	557		94	63	69	783

Within the categories high-falling (H*(+)L) and high-level (H*), the transcription further distinguished between object accents whose peaks stepped *down* relative to the peak of the preceding subject accent (labels !H*L and !H*), and object accent peaks stepping *up* relative to the preceding subject accent (labels ^H*L and ^H*). Table 4.8 presents the results of this sub-classification. For the sake of clarity, H*L and H*+L are collapsed here. As a matter of fact, differentiating the prenuclear and nuclear H*L accents after their type of fall (with H*L reflecting a gradual fall, H*+L a steep fall) did not yield systematic differences between statements and questions.

Table 4.8. Distribution of downstepped and upstepped high-falling and high-level object accents, broken down by utterance type.

	High-falling			High-level		
	H*(+) L	!H*L	^H*L	H*	!H*	^H*
ST	103	42	55	0	0	0
WQ	31	10	48	0	1	2
YQ	27	0	118	12	0	16
DQ	17	1	105	17	0	46
Downstep		53			1	
Upstep			326			64
Total	178	53	326	29	1	64

The number of downstepped accents was greatest in the statements, which is only to be expected if it is true that downstep gives an utterance a certain ring of finality (cf. Ladd 1996:126, ToDI). Nonetheless, WQ also featured some downstepped accents. Theoretically more interesting were the upstepping accents. As the table shows, there were altogether 390 instances (326+64) of this step-up in pitch causing the object accent peak to be perceived as distinctly more salient than the subject accent peak. However,

whether this auditory impression was exclusively brought about by an upstepped object accent or, rather, by a reduced subject accent, or by a combination of these two, has to be determined below on the basis of the f0 measurements. At this point it suffices to say that, at least auditorily, a good many of the questions seemed to feature a distinct asymmetry between subject and object accent. In fact, from Table 4.8 two opposing trends become apparent, suggesting a dichotomy in the four utterance types under investigation. On the one hand, *downstep* mainly occurred in ST and WQ whereas it was virtually absent from YQ and DQ. On the other hand, the incidence of *upstep* was substantially greater in YQ and DQ than in ST and WQ. In Table 4.9, the results are reshuffled to reflect these different tendencies. The column 'equal' refers to the accents which neither stepped down nor up.

Table 4.9. Peak relations between subject and object accent.

	Downstep	Equal	Upstep
ST + WQ	53	134	105
YQ + DQ	1	74	285

Evidently, between the pairs ST/WQ and YQ/DQ there was a distinct decrease in downstep and a distinct increase in upstep. As a matter of fact, in YQ/DQ an upstepped object accent seemed the unmarked choice, considering that it occurred almost four times as often as the equal accent configuration. Accordingly, in YQ/DQ the incidence of 'equal' dropped to almost half of that in ST/WQ. A chi-square test indicated that the proportions of accent types in the two utterance pairs were significantly different ($\chi^2_{(2)}$=150.4.1, p<.001), that is, there was a significant association between utterance-pair and type of object accent (downstepped vs. equal vs. upstepped). Pairwise chi-squares showed that the association between 'equal' and 'downstep' was significant ($\chi^2_{(1)}$ =21.1, p<.001), as was the association between 'equal' and 'upstep' ($\chi^2_{(1)}$ =83.2, p<.001).

4.2.2.2.5 Final boundary tones

As expected, questions predominantly ended with H%; labelling the final boundary tones did not pose any problems. The high boundary tone usually manifested itself as an abrupt upward movement of pitch in the final (unaccented) syllable. By contrast, when no final rise was present pitch gradually lowered. That is, L% was unlike H% in that it did not represent a distinct *movement.* Rather, pitch petered out, pointing to a basic asymmetry between the two final boundary tones[6]. Table 4.10 presents the incidence of high vs. low final boundary tones across the corpus.

[6] This asymmetry would seem to have a physiological correlate. Strik & Boves (1995) found that, in (Dutch) interrogative utterances (which, presumably, ended with a high boundary tone), subglottal pressure is released only *after* the last voiced sound, whereas a release *during* voiced sounds at the end of the utterance goes together with final lowering (i.e. a low final boundary tone).

Table 4.10. Final boundary tones (distribution).

	ST (200)	WQ (186)	YQ (199)	DQ (198)	Total (783)
L%	200	70	8	0	278
H%	0	116	191	198	505

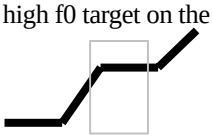
As regards the shapes of the final rises, in the nuclear tones (i.e. the nuclear accent together with the final boundary tone, cf. Grabe 1998:50) four configurations could be distinguished; these are illustrated in Figure 4.1a-d (boxed areas roughly indicate the CV-part of the accented syllable):

(4.1a)

H*L H%

H*L H% ('low H%')
From the high f0 target on the object accent, pitch fell, then rose again.

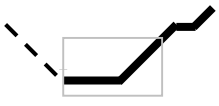
(4.1b)



H* H%

H* H% ('high H%')
Following the high f0 target on the object accent, pitch remained level for a brief stretch; from this high level, the final rise took off.

(4.1c)



L*H H%

L*H H%
At the end of the low-accented syllable pitch immediately rose to form a brief level plateau from which the final rise took off.

(4.1d)

L* H%

L* H%
At the end of the low-accented syllable pitch remained low, to rise again in the final unaccented syllable.

Note that the configurations (4.1b) H*H% and (4.1c) L*H H% were similar in that both maximised the proportion of high pitch. That is, the former altogether dispensed with a low (accentual trailing) tone, and in the latter the low (starred accentual) tone was immediately followed by a high (accentual trailing) tone. The issue of whether these four rise-configurations are somehow related is addressed in chapter 5 (§ 5.3.2.1). For now it suffices to say that, in calculations of mean f0 values, contours (4.1a) and (4.1b) were kept apart whenever this was expedient. Given the low incidence of low pitch accents, the contours (4.1c,d) were excluded from many of the quantitative analyses anyway. Table 4.11 shows the distribution of the four different nuclear tone patterns across the three question types.

Table 4.11. Nuclear tone patterns
Distribution of the four different nuclear tone patterns ending with H%, broken down by question type (N=467, representing the total number of questions in which the object accent was realised and was followed by a high final boundary tone).

	Question type			
Nuclear tone	WQ	YQ	DQ	Total
H*L H%	50 (64%)	137 (72%)	123 (62%)	310 (66%)
H* H%	3 (3.5%)	28 (15%)	63 (32%)	94 (20%)
L*H H%	2 (2.5%)	9 (4.5%)	10 (5%)	21 (5%)
L* H%	23 (30%)	17 (8.5%)	2 (1%)	42 (9%)
Total	78 (100%)	191 (100%)	198 (100%)	467

Apparently, the sequence H*L H% represents the obvious choice; the other three contours can be seen as more idiosyncratic. Even so, the table shows that the three question types differed in the use they made of these less canonical contours. Thus, percentages of contours maximising high pitch (i.e. H* H% and L*H H%) increased in the order WQ<YQ<DQ, whereas percentages of the contour minimising high pitch (i.e. L* H%) showed a decrease. Associations between question type and contour type

(maximalised vs. minimalised high pitch) were significant ($\chi^2_{(2)}$=33.6, p<.001). Pairwise, the associations were also significant: WQ~YQ: $\chi^2_{(1)}$=5.7, p<.017, and YQ~DQ: $\chi^2_{(1)}$=14.8, p<.001. These systematic associations squared with the Functional Hypothesis as well as with the earlier observation that, typically, YQ and DQ pattern together against WQ (and ST), in that the former tend to have more high pitch than the latter.

4.2.2.3 Summary and conclusions

From a comparison of the transcription labels assigned to the statements and the questions in our corpus, the following conclusions can be drawn:

- In both discoursal categories, **initial boundary tones** were labelled predominantly low (statements: 96%, questions: 89%). This went against the *a priori* assumption that at least WQ and YQ might feature large numbers of %H, given that their onsets, unlike those of DQ, proved significantly higher than those of the statements (cf. chapter 3, § 3.5.2.2). However, given our earlier finding that the questions generally displayed a raised register level, %L in YQ could at the same time be phonetically higher than in the statements, yet relatively low within the overall register of the individual utterance. In WQ, the initial syllable was accented, making it more difficult to distinguish boundary tone from accent tone. Yet, in a majority of the cases the boundary tone was labelled %L because the (high-falling or high-level) accent on the wh-word was preceded by a short rise. Also, it was felt that insertion of an unaccented monosyllablic word before the wh-word would most likely have resulted in a *low* boundary tone.

- Except for one case, **wh-words** were always accented; accents were mainly high-falling (71%); often, the fall was steep and large. In a further 28%, pitch following the accented wh-word remained high-level, forming a plateau.

- In the questions, deaccentuation of the **subjects** occurred in 201 out of the 583[7] tokens (i.e. in 35%), predominantly in WQ (i.e. in 88% of the 186 neutral WQ renderings). By contrast, in the 200 statements there was only one case of deaccentuation.

- In ST, YQ and DQ, **objects** were usually accented. In WQ, however, deaccentuation of the object was a frequent occurrence (37%); possibly, this related to the strong gravitational effect of the utterance-initial wh-word.

- In a majority of the questions (74%), pitch between subject accent and object accent was perceived to step up; this occurred mainly in YQ/DQ (i.e. in 63% out of the 74%).

[7] I.e. the number of neutrally realised questions.

- In conjunction with the nuclear accents, **final boundary tones** exhibited four different patterns. Of these, the most commonly occurring was H*L H% (66%). Of the three others, two involved a greater proportion of high pitch and their incidence in the questions followed the hypothesised order WQ<YQ<DQ. That is, they occurred most in the question type that depends entirely on intonation and that, according to the Functional Hypothesis, was expected to feature a maximum of high(er) pitch.

4.2.2.4 Syntagmatic results

4.2.2.4.1 **Statements** (N=200)

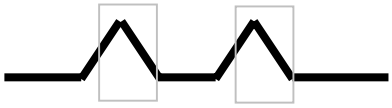
As the present study investigates the intonation of questions against the background of the intonation of their statement versions, we start with a melodic description of the latter. In the statements, two main contour shapes could be distinguished. By far the commonest was the two-peaked contour, pattern 1, which consisted of a low initial boundary tone, two more or less equal accents on subject and object, and a low final boundary tone (transcribed as %L H*L H*L L%; 1a below). This sequence of two so-called 'pointed hats' (cf. 't Hart et al. 1990) occurred in 149 statements (75%) and we may safely regard it as the default pattern. In 48%, the first accent was steeply falling (H*+L), in 52%, it displayed a more gradual fall (H*L); in the second accent, percentages were 68 for the steep fall H*+L, 32 for the more gradual version H*L. In a subset of pattern 1, the second accent gave the auditory impression of being downstepped relative to the first (i.e. %L H*L !H*L L%; 1b). In another subset, pitch between the first and second accent was perceived to step up (%l H*L ^H*L L%; 1c).

In the other (and minor) main pattern (pattern 2), pitch did not drop after the first accent. Rather, retaining its high level it formed a plateau with the peak of the second accent before returning to the baseline. This so-called 'flat hat' (cf. 't Hart *et al.* 1990) occurred in 41 statements (20.5%). Three subtypes could be distinguished, viz. (2a) peak I and peak II were perceived as having roughly the same height (%L H* H*L L%); (2b) peak II was downstepped relative to peak I (%L H* !H*L L%); and (2c) peak II was stepped up relative to peak I (%L H* ^H*L L%).

Figure 4.2 illustrates the patterns and gives percentages of their occurrence. Together, pattern 1 and 2 accounted for 95% of the contours; in the remaining 5%, the first pitch accent was either L*H (N=9) or absent (N=1).

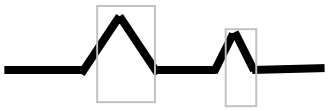
Pattern 1: 75%

(1a: 63%)



%L H*L H*L L%

(1b: 9%)



%L H*L !H*L L%

(1c: 28%)

%L H*L ^H*L L%

Pattern 2: 20%

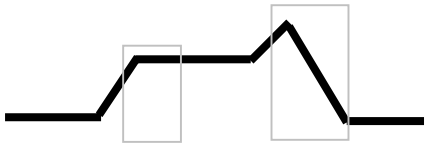
(2a: 17%)

%L H* H*L L%

(2b: 71%)

%L H* !H*L L%

(2c: 12%)



%L H* ^H*L L%

Figure 4.2. Statements: Main accent patterns.

4.2.2.4.2 **Wh-questions** (N=186).

In this question type, also, two distinct patterns could be observed. In pattern 3, the initial wh-word had a short, sharp rise immediately followed by a steep and extensive fall to the baseline (the short rise could be absent, though, cf. § 4.2.2.1 above). In the stretch of low or gradually downsloping pitch following this fall, there were often no further pitch obtrusions, that is, subject and (less frequently) object were not accented (in 88% and 37%, respectively). Pitch then either remained low till the end, or it was followed by a final rise. The sequence, which occurred in 71% of the wh-questions, was transcribed %L H*(+)L L% or %L H*(+)L H%. Figure 4.3 schematises the pattern; dotted lines indicate optionality.

Pattern 3: 71%

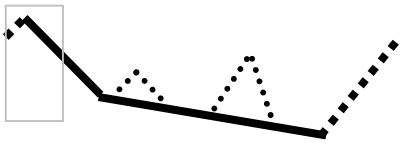


Figure 4.3. Wh-questions: Accent patterns

Pattern 4, likewise, set off from high accentual pitch (which was, similarly, sometimes preceded by a short rise on the question word), but here pitch retained its high level to form a plateau. Following this, there were two possibilities: in variant (4a), pitch made a fall after the plateau (%T H* H*L T%), in variant (4b), prior to the fall, pitch showed a marked step-up (%T H* ^H*L T%). Altogether, patterns with level H* on the wh-word occurred in 29%. Figure 4.4 exemplifies the variants (4a) and (4b).

Pattern 4: 29%

(4a: 14%) (4b: 15%)

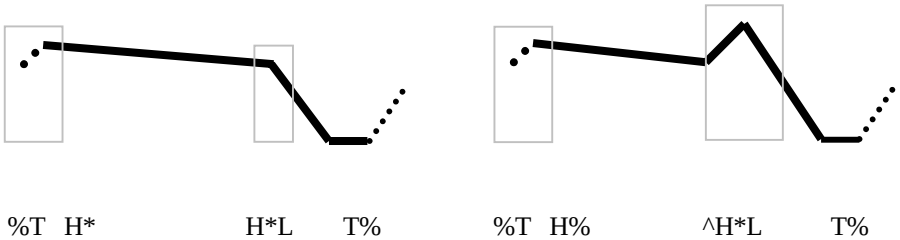


Figure 4.4. Wh-questions: Accent patterns

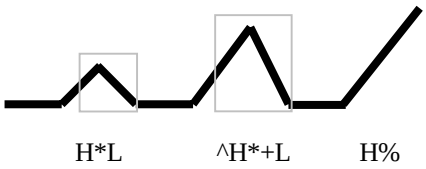
As regards the distribution of the above patterns and variants across speakers, some subjects consistently realised one and the same pattern, that is, either (variants of) 3 or (variants of) 4; others alternated. Apparently, there is some leeway. As regards the final rises (produced in 67% of the wh-questions), the choice for H% vs. L% seemed largely a matter of speaker-specific preference. Out of the four speakers who consistently produced the same final boundary tone, one always preferred L%, the other three H%. The remaining six speakers displayed varying proportions of H% and L% in what seemed a haphazard way. This was in line with findings of Hirschberg's (1989) that the final rises in American-English wh-questions typically showed speaker-variability. Altogether, it was our impression that, in our experimental wh-questions, the presence or absence of a final rise did not substantially affect the interrogative character of the utterance: the wh-questions sounded 'questioning' also without H%. This apparent redundancy of H% is not surprising, of course, considering that WQ featured other cues to interrogativity, viz. the question word (lexical cue), the accent on the question word (intonational cue), and inversion (syntactic cue). This creates scope for the final rise to be used for other, notably pragmatic purposes, possibly in combination with further variables such as context. Thus, Cauldwell (1999) found that the English wh-question *What do you mean?* with a falling last tone on *mean* was judged as 'irritated, angry' when presented in isolation. However, when a context was supplied, the same utterance was rated as *not* conveying anger or irritation. Although our own material provided a context, speakers might still have wanted to produce a final rise on WQ to avoid creating an impression of, say, domination or some other negative attitude that may be associated with a falling WQ. Experiments in this interesting area will have to be left for future research.

4.2.2.4.3 **Yes-no questions** (N=199).

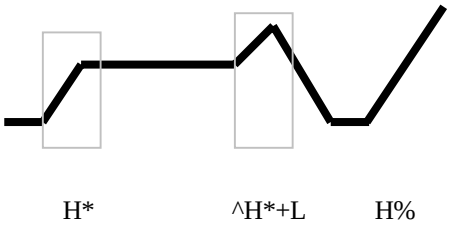
The most striking feature in this question type was a distinct step-up in pitch between the subject accent and the object accent (see Table 4.7 above). On the whole, this asymmetry seemed to be brought about by pitch excursions on the subject accents being relatively small (as compared with those in the statements), whereas object accents seemed at least as large but often larger than their counterparts in the statements. The pattern was labelled %L H*L ^H*L H% (5a below); it occurred in 42% of the cases. Mostly, the high-falling subject accents fell gradually (H*L), whereas the high-falling object accents usually displayed steep falls (H*+L). The step-up in pitch could also be observed in contours where either the *first* accent failed to return to the baseline (the 'flat hat', i.e. H* ^H*L H%; 5b below, 11%), or the *second* ('high H%', i.e. H*L ^H* H%; 5c below, 6%). Together, the three upstepping patterns accounted for 59% of the contours. Of the remaining contours, 16.5% lacked a subject accent, 12.5% had a L*(H) accent, and 12% did not show upstep between the accents. Figure 4.5 gives schematic representations of the upstepping patterns.

Pattern 5: 59%

(5a: 42%)



(5b: 11%)



(5c: 6%)

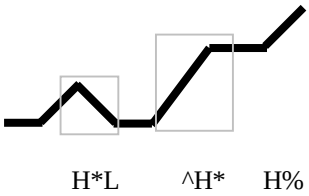


Figure 4.5. Yes-no questions: Upstepping accent patterns

4.2.2.4.4 **Declarative questions** (N=198).

Overall, the intonation patterns in this question type appeared fairly similar to those in the yes-no questions. Thus, of the high-falling subject accents, the majority was gradual, whereas the high-falling object accents mostly featured steep falls. There was, however, a far greater frequency of upstepping accent patterns. Together, these occurred in 76% of the contours, that is, the above pattern (5a) could be observed in 38%, (5b) in 15% and (5c) in 23%. The remaining 24% consisted of contours with L*(H) pitch accents (5.5%), contours without a subject accent (3.5%), and contours without upstep (15%). Table 4.12 compares the patterns across the two question types.

Table 4.12. Percentages of upstepping accent patterns in the yes-no questions and declarative questions

	Upstepping patterns			Total	Residual		
	H*L ^H*L	H* ^H*L	H*L ^H*		no upstep	no 1st acc	L*(H)
YQ	42%	11%	6%	59%	12%	16%	12%
DQ	38%	15%	23%	76%	15%	3.5%	5.5%

In spite of the qualitative similarity between the YQ and DQ contours, there was a suggestion of DQ having somewhat larger subject accents, object accents and final rises. These quantitative details will be discussed in § 4.3 below.

4.2.2.4.5 Incidence of low accents

A few brief comments may be in order with respect to the relatively limited incidence of low pitch accents. Altogether, there were 73 occurrences of this accent type, viz. 30 cases of L*H and 43 of L*; most of these occurred on the object accent (i.e. 63). For a start, considering that 62% of the low accents were produced by the same speaker, this accent type would not seem a very obvious choice. Moreover, the choice for a low accent would seem to have been largely inspired by syntagmatic considerations. Thus, while H*(L) accents could occur after high, high-falling and low initial boundary tones, L*(H) accents were observed only in combination with the former two. Likewise, while H*(L) accents could be found after high as well as low accents, L*(H) accents only followed high accents. And, finally, while H*(L) accents could precede low as well as high final boundary tones, L*(H) accents were seen to precede only H%. These distributional constraints suggest that a low accent has to rely on a neighbouring high tone for it to be effective. This raises some doubts as to the status of the L*(H) pitch accent. Can it be viewed as an accent in its own right, that is, as a fully equivalent counterpart to the H*(L) accent (expressing some abstract meaning of its own, cf. Gussenhoven 1984a, 1988a)? Or, rather, is it primarily motivated by the choice of neighbouring tones, or is it largely speaker-specific?

To be sure, low pitch accents appear to pose more problems for modelling than do their high counterparts. This is apparent in, e.g., Grabe's (1998) comparison of speech corpora of Northern Standard German and Southern Standard British English. Her English corpus yielded very few instances of nuclear L*+H and no instances at all of non-final L*+H. In the German corpus, the shapes of L*+H were found to be considerably less stable than those of H*+L, as were the beginnings of their rises. In fact, apart from being a relatively rare occurrence the German low accent also sometimes looked "strikingly similar to H*+L and [it] cannot be distinguished unless the alignment with the stressed syllable is carefully observed" (p.87). If low accents tend to be little used in speech and if, as in Grabe's study, "a considerable amount of detail about the alignment of f0 with segmental material [is] needed" (p. 88) to label them as instances of L*+H, more research into production and distribution of this pitch accent would seem to be called for (for Dutch distribution data, see chapter 5, § 5.3.2: Caspers 2000b).

4.2.2.5 Summary and conclusions

From the syntagmatic comparison of the statement and question contours, several qualitative differences emerged. A striking feature in the questions was the frequent step-up in pitch between subject and object accent as it occurred in, notably, YQ (59%) and DQ (76%). This phenomenon could be observed both in the high-falling and in the high-level object accents. In the statements, it was found in only 27% of the cases. The second part of the chapter investigates whether the *amount* of upstep was, likewise,

systematically smaller in ST than in YQ and DQ. As regards the accent patterns of WQ, comparison with the other three utterance types was somewhat hampered by an intrinsic asymmetry: in WQ, the subject and object accent slots were preceded by an additional accent slot for the wh-word. This initial accent usually boasted a relatively large excursion, whereas subject and object frequently went without tonal movement (in 88% and 37%, respectively, of the 186 neutral realisations). If there *were* accents on subjects and objects in WQ, their excursions gave the impression of being reduced relative to those in the statements. This perceived dominance of the wh-word was not surprising considering that, pragmatically, the wh-word is the focal point of the utterance. In all, the tonal centre of gravity of WQ seemed to reside in the initial half of the utterance, giving rise to a clear difference with the other two question types (for a quantitative evaluation of this difference, see further below). Although upstep in WQ occurred in 27%, it made the impression of being communicatively less effective, considering that the pitch peaks of subject and object often lay well below that of the wh-word.

Another observation was that the questions featured a relatively large number of high pitch plateaus (label H*). In the questions, such high level stretches occurred in 32%, against 20% in the statements. In WQ, these plateaus directly linked the peak of the question word to the peak of subject or, mostly, object accent. Occurring early in the question, such plateaus further contributed to the acoustic dominance of the initial portion, even when a final rise followed. By contrast, in YQ and DQ plateaus were always positioned *after* the first accent, that is, they either joined the subject accent to the object accent, or the object accent to the final high boundary tone (creating 'high H%'). That is, in YQ and DQ upsteps and high plateaus caused the acoustic emphasis to lie in the *latter* half of the utterance. On the basis of the observations so far, we suggest that the three question types can be roughly divided into two subgroups, one in which high pitch concentrates in the first half of the question (WQ), as opposed to one where this is true of the second half (YQ and DQ). However, further evidence for such a division has to come from the quantitative analyses below.

4.3 Quantitative aspects

4.3.1 Introduction

Following the investigation of qualitative aspects of accent in the corpus of read speech, the second part of this chapter concentrates on quantitative properties. In chapter 3, the testing of hypothesis (iv) was deferred till the present chapter, which explicitly deals with matters of accentuation[8]. Hypothesis (iv) predicted that, in questions, the (nuclear) accent is raised. This would mean that, in our two-accented Dutch question material, the accent on the object is raised relative to the accent on the subject. An indication that this is indeed the case already came from the frequent assignment of the transcription label 'upstepped' to the nuclear accents in, notably, the yes-no and declarative questions (cf. § 4.2.2.2.4 above). Earlier, in chapter 3, the upward slopes of the *upper* regression lines in YQ and DQ (Figure 3.12) also suggested that the latter halves of these question types featured relatively much high pitch (note that the regression lines did not include the

[8] The hypotheses (i), (ii), (iii) and (v) were tested in chapter 3.

final rises). However, we need to make a quantitative assessment of the upstep phenomenon, and in the following sections, this issue will be addressed. In § 4.3.2, we motivate the choice of parameters. Next, § 4.3.3 deals with the acoustic and auditory analyses of the corpus material, that is, § 4.3.3.1 describes the method and measurements and § 4.3.3.2 presents the results. In § 4.4, the conclusions are summarised, § 4.5 discusses the main results, and § 4.6 provides a general summary.

4.3.2 Parameters

It was not immediately obvious how hypothesis (iv) should be operationalised. Issues involving peak relations may be complicated by, among other things, a discrepancy between production and perception. Experiments with two-accent statements have shown that a second accent peak has to be *produced* on an objectively *lower* f0 value than the first for it to be *perceived* as being of *equal* prominence (e.g. English: Pierrehumbert 1979; Dutch: Rietveld & Gussenhoven 1985; Gussenhoven & Rietveld 1988:358; Gussenhoven *et al.* 1997). This raised the question as to when a second accent peak can be considered 'raised'. However, as the present study rests on speech *production*, it seemed appropriate to take the production perspective. Hence, a second accent peak (i.e. on the object) was regarded as 'raised' as soon as its f0 value exceeded that of the preceding peak (i.e. on the subject). Possible systematic differences between questions and statements were expected to show up in comparisons of the inter-peak intervals across the four utterance types. Accordingly, the first parameter, Δ PEAKS, involved the difference in f0 between peak II and peak I, with a positive value indicating a raised nuclear accent peak.

However, a raised accent may also manifest itself as a larger excursion size. That is, while the peak f0 of a second accent may be lower than that of its predecessor, its excursion size may actually be greater due to a gradual lowering of the baseline[9] from which the second accent takes off. Therefore, the second parameter, Δ EXCURSIONS, represented the difference in excursion size between accent II and accent I with, again, a positive value pointing to a raised second accent. Obviously, calculation of these two syntagmatic measures presupposed paradigmatic measures, viz. the individual peak heights of subject accent and object accent, as well as the individual excursion sizes. These were included as parameters in their own right, together with the values of the peaks and excursions of the wh-word. This added up to a set of eight parameters:

- Peak f0 of
  - (i) wh-word;
  - (ii) subject accent;
  - (iii) object accent;

[9] Since there is, as yet, no conclusive evidence as to how accents are being scaled and how this scaling affects perception, the term 'baseline' may refer to several notions, i.e. an observable line drawn through a contour's valleys, or some more abstract line of reference (e.g. Pierrehumbert 1980, Liberman & Pierrehumbert 1985, Ladd 1993, Gussenhoven *et al.*1997).

- Excursion size of
  - (iv) wh-word;
  - (v) subject accent;
  - (vi) object accent;

- (vii) Δ Peaks (f0 object peak – f0 subject peak)

- (viii) Δ Excursions (exc. object peak – exc. subject peak)

With respect to, specifically, the syntagmatic parameters Δ PEAKS and Δ EXCURSIONS we were interested to establish which of the two showed the strongest effect across speakers and utterance types. That is, is it the relation between *peak* values that is the most carefully controlled by speakers, or the relation between *excursion* values? In the next section, this issue is further elaborated on.

4.3.2.1 Accent relations and focus structure

In a classic production experiment, Liberman & Pierrehumbert (1984:210) carried out a quantitative investigation of the accents in the two-accent statement *Anna goes with Manny*, in which they had systematically varied the focus structure. Thus, either *Anna* constituted new information and the rest 'background', or it was the other way round, with *Manny* constituting the new information, *Anna* the old. The accent patterns corresponding to these two focus structures were elicited by making the utterances appropriately-sounding answers to the questions 'What about Manny? Who came with him?' vs. 'What about Anna? Who did she come with?'. Given that 'new information' was signalled by means of a relatively large accent, and 'background' by means of a relatively small one, two accent patterns emerged. One pattern showed a distinct step up in pitch between the accents on *Anna* and *Manny* (similar to our questions), the other a distinct step down. Four subjects realised these two patterns in ten different pitch ranges, varying between a 'phlegmatic mumble' and a 'forceful shout'. When Liberman & Pierrehumbert compared the two accent pairs in terms of absolute peak levels and of excursion sizes they found, among other things, that there was a constant ratio between the peak levels, irrespective of the decrease or increase in range. This suggested that the peak relations had been under careful control of the speakers, a control that seemed absent from the relations between the accentual rises, or between the accentual falls, i.e. between excursions.

A perception experiment in this area was conducted by Rump & Collier (1996). They presented listeners with the Dutch two-accent utterance *Amanda goes to Malta* and asked them to adjust the heights of either accent peak relative to the other (which was fixed). By manipulating the accent relations, the subjects were expected to correctly reproduce four different focus structures: neutral focus (answering the question 'What is happening?'), contrastive focus on *Amanda* (question: 'Is John going to Malta?'), contrastive focus on *Malta* (question: 'Is Amanda going to Cyprus?') and double contrastive focus (question: 'Is John going to Cyprus?'). On the basis of these adjustments, prototypical pitch contours of each focus structure were set up (cf. chapter

6, Figure 6.2). When, in a follow-up experiment, new listeners were asked to identify these focus structures, this proved an easy task and the results were clear-cut. Rump & Collier's experiments showed that "listeners have clear ideas about how peak heights signal different focus structures." The authors concluded: "[…] appropriate values of the pitch peaks are present in the listeners' (and speakers') minds, serving as a kind of target value which should be reached in the case of prototypical pitch contours." (p. 14). What both the Anna/Manny and the Amanda/Malta experiments indicated is (i) that a particular relation between successive accent peaks corresponds with a particular focus structure, and (ii) that for different focus structures, speakers have internalised certain target values for the accent peaks.

The auditory differences between statements and questions recorded by the transcription of our corpus material showed that, in the questions, pitch generally displayed a clear step-up between the subject and object accent. However, before we can say whether or not these accent relations reflected a certain focus structure, we have to establish whether the perceived differences correspond to statistically significant differences in f0 values. This will be undertaken below.

4.3.3 Acoustic/auditory analysis

4.3.3.1 Method and measurements

In each of the 800 utterances, the pitch configurations on wh-word, subject and object were interactively stylised as rise-fall sequences (or, in the relatively infrequent cases of a low accent, as fall-rises). Next, f0 was measured at the accents' pivot points, i.e. at the low beginning of a H*L accent, at its peak, and at the lowest point following the peak (cf. Liberman & Pierrehumbert 1984). Obviously, with L*H accents it was the reverse. Figure 4.6 shows the relevant pivot points (maximally nine).

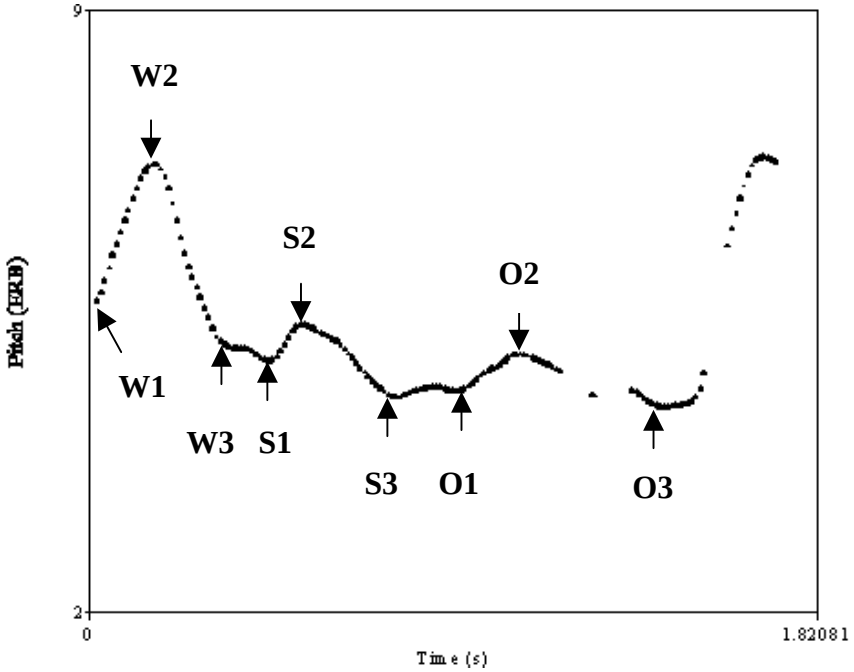


Figure 4.6. Measured pivot points in accents on the wh-word (W1, W2, W3), the subject (S1, S2, S3) and the object (O1, O2, O3).

Frequently, a fall extended beyond the accented syllable. In such cases, its lowest point was measured on the assumption that it is the full size of a fall that is perceptually relevant rather than just part of it. However, stretches representing final lowering were explicitly excluded from the measurements of the accentual fall. Usually, there was an obvious transition between the (steep) fall of the nuclear accent and the remainder of the utterance, that is, in the absence of a final rise pitch mostly levelled off before petering out altogether. Sometimes, pitch accents merely consisted of a rise-cum-peak or a peak-cum-fall. This happened in high plateaus, i.e. whenever accents were linked together (‘flat hat’), or when the nuclear accent peak was directly linked to the final rise (‘high H%’). In such cases, only two of the three pivot points could be included in the data set. To establish the excursion sizes of accents featuring both a rise and a fall, the largest of these two movements was taken into account, again on the assumption that it is the *size* of a movement that carries the greatest perceptual weight.

On these data, eight analyses of variance (SPSS Manova, repeated measures design) were carried out. The choice for separate analyses rather than an overall analysis including all dependent variables was motivated earlier in chapter 3 (§ 3.5). In point of fact, the very circumstance that WQ had three accents instead of two made it impossible for the accentual data to be lumped together in a single analysis. Of the eight dependent variables, three concerned raw values, i.e. WH-

PEAK, SUBJECT PEAK and OBJECT PEAK[10]. The remaining five were derived, that is, WH-EXCURSION, SUBJECT EXCURSION and OBJECT EXCURSION were calculated by subtracting the lowest f0 in the accent from the peak f0. The measures PEAK DIFFERENCE (Δ PEAKS) and EXCURSION DIFFERENCE (Δ EXCURSION) were obtained by subtracting the peak value or excursion size of the subject accent from the peak value or excursion size of the object accent (within the same utterance). Thus, positive Δ-values indicated that object peaks were higher or that object excursions were larger than their counterparts in the subject accents; negative Δ-values reflected the reverse.

As in chapter 3, the analyses of variance had one between-subject factor, SEX OF SPEAKER (two levels, male and female), and three fixed within-subject factors: UTTERANCE TYPE (four levels: statement, wh-question, yes-no question, declarative question), UTTERANCE LENGTH (two levels: short and long), and UTTERANCE POSITION (three levels: in isolation, in first or in second position within a pair). The ten speakers were nested within SEX OF SPEAKER as a random factor. In view of the large number of analyses the level of significance was, again, fixed at $p<.01$ (Huynh-Feldt corrected). As *p* values were not always comparable due to differences in degrees of freedom, we report correlation ratios (η^2) indicating the effect strength (cf. Rietveld & Van Hout 1993:59; see also chapter 3, § 3.5). Only significant effects and significant interactions are reported. For post hoc comparisons, Tukey tests were used; here the significance level was set at .05. Missing values were treated as described in chapter 3, § 3.5. It must be added that, notably in WQ, the problem of empty cells returned with a vengeance. As shown in the Tables 4.6 and 4.7, in the 186 tokens of WQ only 23 subjects and 117 objects were accented, and in the 199 tokens of YQ, 32 subjects were deaccented. This was bound to result in empty cells, given that each of the four utterance types was split up in as many as six conditions (two lengths x three positions, see chapter 3, Figure 3.4). Also, in WQ two of the male speakers never realised an accent on the subject at all. This meant that for these speakers, in analyses involving the subject accents, all six cells referring to WQ-conditions were empty. Since we did not want SPSS to discard all further data of these speakers in one fell swoop[11], we filled the empty cells with the mean values of the other three male subjects. Admittedly, this procedure caused the subsequent statistical analyses to be slightly flawed, at least where WQ was concerned. At the same time, however, it demonstrated that there was a limit to what the statistical package in question could achieve, and there was nothing for it but to accept it. For an optimal interpretation of the results, the reader is also referred to the tables and figures presenting the mean values for each utterance type.

[10] Note that means of peak values and excursion sizes always refer to the H*(L) accents; the relatively few L*(H) accents were left outside consideration.

[11] I.e., their data in the six conditions in ST, in YQ and in DQ (18 conditions).

4.3.3.2 Results

The paragraphs below indicate for each parameter whether or not the results of the statistical analyses reflected systematic accentual differences between the utterance types. Given that, for anatomical reasons, the female voice is inherently higher than the male voice, significant effects of the factor SEX OF SPEAKER on the *raw* parameters are mentioned only in Table 4.18, which summarises the statistical results; elsewhere, they are ignored.

4.3.3.2.1 Accents on wh-words: Peaks and Excursions.

As wh-words occurred only in the wh-questions, there could be no effect of the factor UTTERANCE TYPE here. F0 of accent peaks on the wh-words *Wat?* and *Waar?* did not differ significantly as a function of UTTERANCE LENGTH or UTTERANCE POSITION; the same held for the mean excursion size of the wh-word. This latter (non-raw) parameter did not show significant effects of the factor SEX OF SPEAKER, either.

Figure 4.7 shows that there was a clear tendency for both sexes to realise the wh-peak higher than the subsequent peaks on subject and object. This relative prominence of the wh-word confirmed the auditory impression that subject and object accents were intonationally subordinate to the accent on the wh-word. Presumably, this reflects that the referent of the wh-word contains the specific information sought by the speaker.

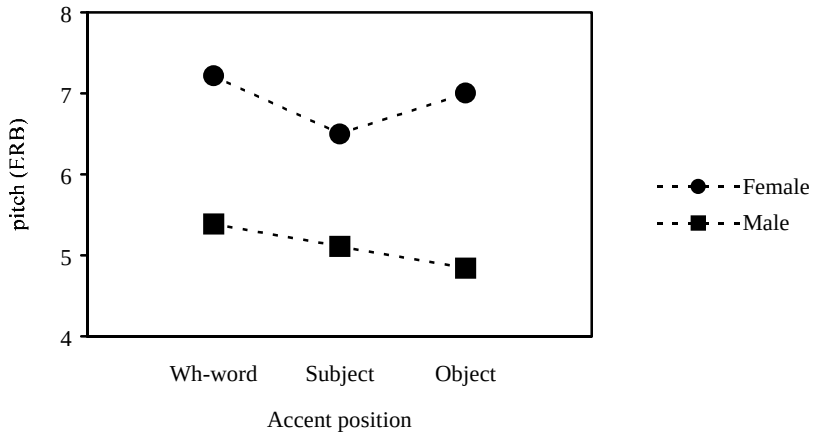


Figure 4.7. Wh-questions: Accent peaks
Mean f0 (ERB) of accent peaks on wh-words (N=185), subject accents (N=23) and object accents (N=92), broken down by SEX OF SPEAKER.

4.3.3.2.2Subject accents: Peaks

The peak levels of the subject accents were not found to differ significantly as a function of the factor UTTERANCE TYPE. As Figure 4.8 shows, in the male speakers subject peaks were roughly equal in all four utterance types. In the female speakers, the peaks of the three question types were somewhat lower than in ST. However, no significant interaction was found between SEX OF SPEAKER and UTTERANCE TYPE.

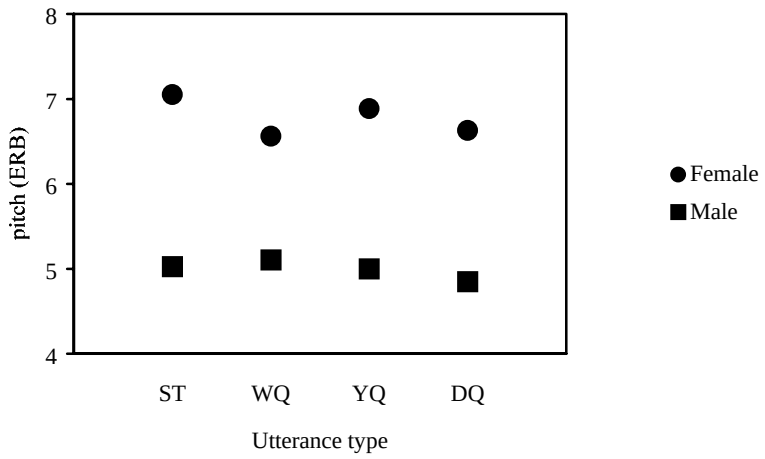


Figure 4.8. Subject peaks
Mean f0 (ERB) of subject peaks in the statements (N=190), wh-questions (N=23), yes-no questions (N=143) and declarative questions (N=180), broken down by SEX OF SPEAKER.

If, earlier, the transcription suggested a step-up in pitch between subject accent and object accent in, notably, YQ and DQ (as opposed to ST), this was not corroborated by significantly lower subject peaks. However, this would not seem hard to explain, considering that all three question types were realised on a significantly higher register than were the statements (see chapter 3). Thus, if the subject peaks in YQ and DQ were not significantly lower than those in ST, this may have been due to an upward shift of the valleys from which the subject accents were scaled. This interaction would seem to present an obstacle to a straightforward comparison of absolute peak values across utterance types whose register levels can be supposed to vary. Hence, there is all the more reason for looking carefully into the results on the subjects' *excursion* sizes.

4.3.3.2.3Subject accents: Excursions

The effect of the factor UTTERANCE TYPE on the excursion sizes of the subject accents was highly significant. The value for η^2 (expressing the effect strength) was quite high, too. Table 4.13 gives the statistical results, together with the post hoc comparisons. Figure 4.9 visualises the results; for the representation of pitch intervals, we use bar graphs.

Table 4.13. Subject Excursions
Statistical results factor UTTERANCE TYPE

Overall effect			Pairwise comparisons					
$F_{(3,24)}$	p	η^2	ST~ WQ	ST~ YQ	ST~ DQ	WQ~ YQ	WQ~ DQ	YQ~ DQ
38.44	.000	.833	.000	.000	.006		.009	.033

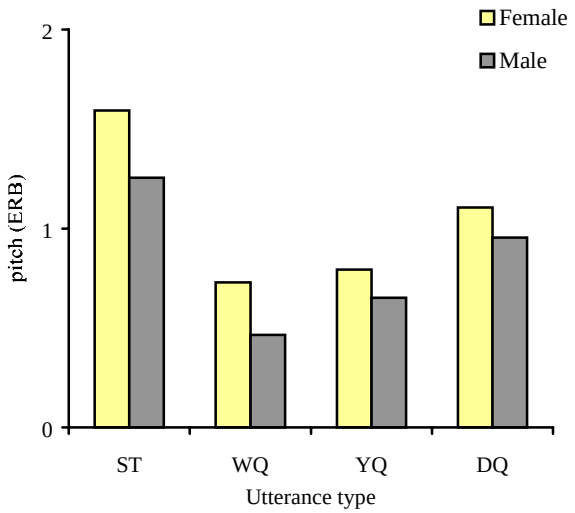


Figure 4.9. Subject excursions
Mean excursion sizes (ERB) of subject accents in the statements (N=190), wh-questions (N=23), yes-no questions (N=143) and declarative questions (N=180), broken down by SEX OF SPEAKER.

Evidently, in each of the three question types, the excursions were smaller than in the statements[12]. Also, the patterns of the female and male speakers were remarkably alike. Although it can be seen that the excursions of the female speakers were always larger, this difference fell short of significance. The difference between ST and DQ was the smallest; it may be recalled that, earlier, another quantitative similarity between these two was observed, viz. in the f0 levels of their onsets (cf. chapter 3, § 3.5.2).

The factor UTTERANCE LENGTH also had a significant main effect, in that the mean subject excursions were larger in the short condition (i.e., when there were only two syllables between subject and object accent rather than five): $F_{(1,8)}=21.89$, $p<.002$, $\eta^2=.732$). Also, there were two significant two-way interactions, viz. (i) UTTERANCE LENGTH x UTTERANCE POSITION, to the effect that in the short utterances subject excursions increased in the order isolation < 1st position < 2nd position, whereas in the long utterances this order was reversed ($F_{(2,16)}=17.84$, $p<.000$, $\eta^2=.690$), and (ii) SEX OF SPEAKER x UTTERANCE POSITION, ($F_{(2,16)}=6,75$, $p<.008$, $\eta^2=.458$), to the effect that in the women subject excursions decreased in the order isolation<1st position<2nd position, whereas the men displayed the reverse order. As we did not have specific hypotheses about subject excursion vis-à-vis UTTERANCE LENGTH, UTTERANCE POSITION or SEX OF SPEAKER, we do not further pursue this part of the results.

To sum up, the two discourse categories statement and question could be fairly easily distinguished on the basis of the excursion size of the subject accent which, in the questions, was systematically reduced.

4.3.3.2.4Object accents: Peaks

The factor UTTERANCE TYPE also had a significant effect on the peak f0 of the object accents. This is shown in Table 4.14 and Figure 4.10.

Table 4.14. Object Peaks
Statistical results factor UTTERANCE TYPE

Overall effect			Pairwise comparisons					
$F_{(3,24)}$	p	η^2	ST~ WQ	ST~ YQ	ST~ DQ	WQ~ YQ	WQ~ DQ	YQ~ DQ
13.13	.000	.621		.010	.003	.001	.000	

[12] Means (in ERB) were: ST 1.41, WQ .65, YQ .71 and DQ: 1.03.

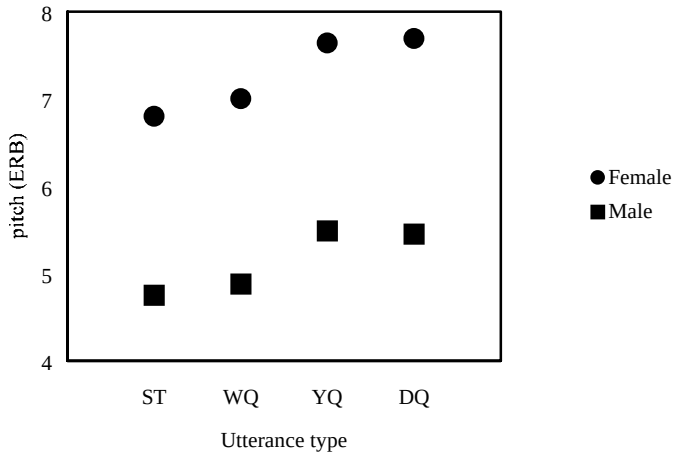


Figure 4.10. Object peaks
Mean f0 (ERB) of object peaks in the statements (N=200), wh-questions (N=92), yes-no questions (N=173) and declarative questions (N=186), broken down by SEX OF SPEAKER.

Clearly, in YQ and DQ, the object accents featured higher peaks than in the statements. This may have been due to their being scaled from a higher register level, as well as to their being locally raised. At present, it would not seem possible to determine the relative contributions of these factors. As with the subject excursions, male and female patterns were strikingly similar.

4.3.3.2.5 Object accents: Excursions

As Table 4.15 shows, the effect of the factor UTTERANCE TYPE on the object excursions fell short of the pre-set significance level of .01. Figure 4.11 illustrates the observed differences.

Table 4. 15. Object Excursions
Statistical results factor UTTERANCE TYPE

Overall effect			Pairwise comparisons					
$F_{(3,24)}$	p	η^2	ST~ WQ	ST~ YQ	ST~ DQ	WQ~ YQ	WQ~ DQ	YQ~ DQ
(5.78)	(.019)	(.420)				.034	.026	

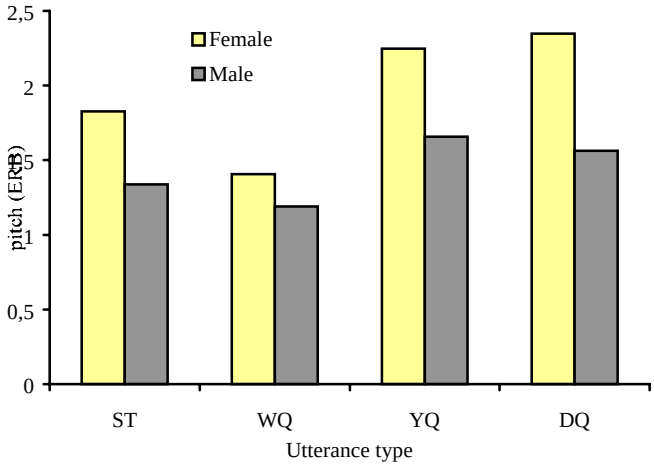


Figure 4.11. Object excursions
Mean excursion sizes (ERB) of object accents in the statements (N=200), wh-questions (n=92), yes-no questions (N=173) and declarative questions (N=186), broken down by SEX OF SPEAKER.

What is of interest, however, is that by now we seem to be able to descry the upstepping mechanism suggested by the transcription of the material (cf. § 4.2.2.2). While YQ and DQ had *subject excursions* that were significantly *smaller* than those in the statements, their *object peaks* were at the same time significantly *higher* (cf. Figures 4.9 vs. 4.10). (There is no such effect in WQ, whose subject and object excursions were both relatively small (smaller than in ST), and whose mean subject and object peaks lay below the mean peak on the wh-word (cf. Figure 4.7)). However, it should be borne in mind that the results so far are paradigmatic, they do not reflect utterance-*internal* relationships between subject and object; these are expected to emerge from the Δ parameters.

At this point, a remark about the subject and object accents occurring in WQ may be in order. It was seen above that the former were deaccented in 88% of the 186 utterances, the latter in 37% (cf. § 4.2.2.4.2). Moreover, the Figures 4.9 and 4.11 showed that the accents in WQ that *were* actually realised had the smallest excursion sizes of all four utterance types. This probably explains why this question type featured a narrowed register span (see chapter 3, Figure 3.14). When the regression lines were calculated, WQ either had no accent data points at all because of deaccentuation, or the excursions of these accents were relatively small.

4.3.3.2.6 Δ Peaks

Naturally, results on the factor Δ PEAKS indirectly reflected the individual results on SUBJECT PEAK and OBJECT PEAK, however without the disadvantage of raw values that could not be properly compared across the sexes (and across register levels). As Table 4.16 illustrates, the intervals between object peak f0 and subject peak f0 were

significantly affected by the factor UTTERANCE TYPE. Figure 4.12 shows the direction of the differences.

Table 4. 16. Δ Peaks
Statistical results factor UTTERANCE TYPE

Overall effect			Pairwise comparisons					
$F_{(3,24)}$	p	η^2	ST~ WQ	ST~ YQ	ST~ DQ	WQ~ YQ	WQ~ DQ	YQ~ DQ
13.19	.000	.622		.003	.000	.034	.004	

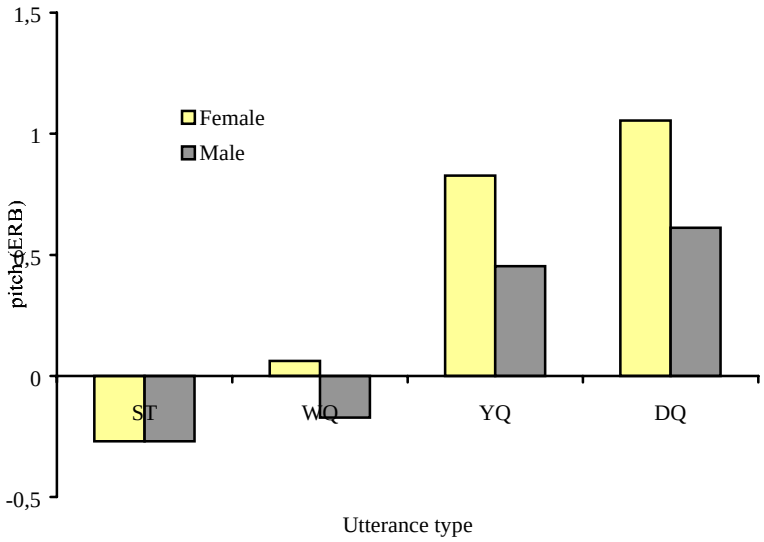


Figure 4.12. Δ Peaks
Differences (in ERB) between the f0 of object peaks and subject peaks (Δ obj–subj), broken down by UTTERANCE TYPE and SEX OF SPEAKER

What Figure 4.12 now effectively brings to light is the presence vs. absence of the upstepping mechanism between the subject and object accents. In ST, this phenomenon was clearly absent: since the object peak was lower than the subject peak, values on this parameter were negative. In contrast, YQ and DQ displayed positive values; in DQ, the amount of upstep was evidently largest. As for WQ, here the differences between the two peaks were even smaller than in ST. Also, the male and female patterns slightly diverged here. Consequently, the women produced upstepping pitch in all three question types. Moreover, their amount of upstep in YQ

and DQ was distinctly greater than in the men (without reaching significance, though).

A significant interaction was found between UTTERANCE TYPE and UTTERANCE POSITION, to the effect that in YQ and DQ the inter-peak interval was largest in utterances in first position within the pair, whereas in ST and WQ, this was not the case: $F_{(6,48)}=4.43$, $p<.005$, $\eta^2=.357$. Considering that we did not have hypotheses about possible relationships between these two parameters, we will not attempt to offer an explanation for their interaction.

4.3.3.2.7 Δ Excursions

The parameter Δ EXCURSIONS reflected the differences between the sizes of subject and object excursions, themselves derived parameters. As can be seen in Table 4.17, there was a significant main effect of the factor UTTERANCE TYPE, but this was weaker than in Δ PEAKS. Also, among the utterance types only the difference ST~YQ reached significance. Figure 4.13 gives a graphic representation.

Table 4. 17. Δ Excursions
Statistical results factor Utterance Type

Overall effect			Pairwise comparisons					
$F_{(3,24)}$	P	η^2	ST~ WQ	ST~ YQ	ST~ DQ	WQ~ YQ	WQ~ DQ	YQ~ DQ
7.00	.003	.467		.003				

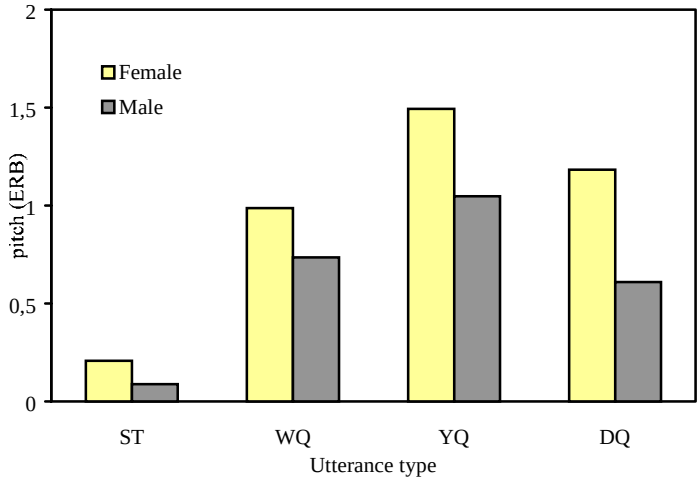


Figure 4.13. Δ Excursions
Differences (ERB) between the excursion sizes of object accents and subject accents, broken down by UTTERANCE TYPE and SEX OF SPEAKER.

There were no negative values which meant that, in all four utterance types, the object excursions were larger than the subject excursions; in the statements, this difference was very small, though.

Note that the lower values of DQ on Δ EXCURSIONS (Figure 4.13) reflect the earlier finding that, in this question type, the excursion sizes of the subjects were less reduced relative to those in the statements than they were in WQ and YQ (cf. Figure 4.9).

4.3.3.2.8Sex of Speaker

With respect to their accent realisations across the four utterance types, the two sexes displayed broadly similar patterns. Where the parameters SUBJECT EXCURSION, OBJECT PEAK and Δ EXCURSIONS were concerned, this similarity was even striking. The few minor divergences mainly concerned WQ, the question type where the low number of realised accents rendered the results somewhat less stable.

Statistically, the factor SEX OF SPEAKER was found to have significant effects only on the variables expressed in raw values, viz. WH-PEAK, SUBJECT PEAK and OBJECT PEAK. We may assume that this part of the results was mainly or solely due to the intrinsic pitch differences between men and women. However, Figures 4.9, 4.11, 4.12 and 4.13 make it clear that, on the whole, the values of the female speakers on the relevant derived parameters exceeded those of their male counterparts, notably in YQ and DQ. Thus, the women produced larger excursions on their subject and object accents, a greater upstep between subject and object accent, and a greater difference between the excursion sizes of subject and object accent[13]. Figure 4.14 sums up this trend by comparing the men and women on their performance on the (derived) parameters EXCURSION SUBJECT and EXCURSION OBJECT (and, implicitly, on Δ EXCURSIONS).

[13] Recall that all f0 values are expressed in ERB, a scale that is assumed to allow direct comparison of male and female speech.

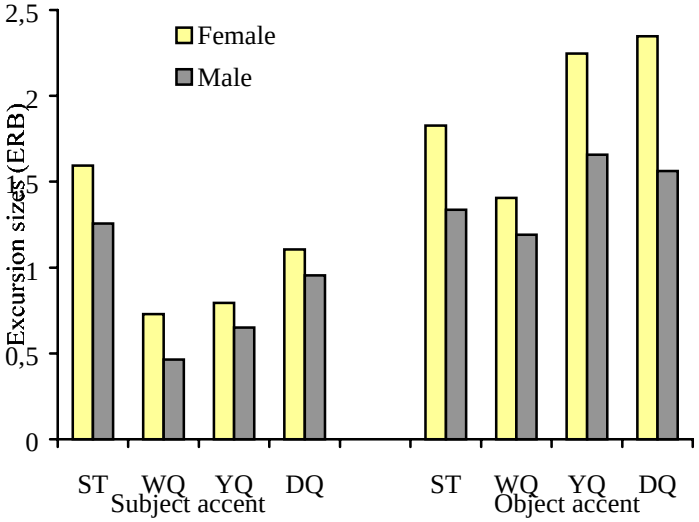


Figure 4.14. Excursion sizes as a function of SEX OF SPEAKER.

Although the differences between the sexes did not reach significance, they were systematic enough to speak of a tendency in the female speakers to locally expand their accents. This met the expectation (formulated in chapter 3, § 3.3.6.2) that sex-based differences in register span may show up in *local* pitch movements. It is also in line with Patterson & Ladd's (1999) claim, that a (linguistically motivated) measure of pitch range based on excursion sizes of accents is more successful than a measure based on an overall maximum-minimum difference.

To conclude the sections on the quantitative results, Table 4.18 presents an overview of the main effects and Figure 4.15 gives prototypical accent patterns for each of the four utterance types under investigation, produced by the same (male) speaker.

Table 4.18. Main effects.
Significance levels (.01) and η^2 of the four factors SEX OF SPEAKER, UTTERANCE TYPE, UTTERANCE LENGTH, and UTTERANCE POSITION. Between brackets: significances at the .05 level.

Factors	Sex		Type		Length		Position	
	p	η^2	p	η^2	p	η^2	p	η^2
Wh-peak	.001	.797	-	-	(.044)	.417	(.033)	.361
Wh-excursion	-	-	-	-	-	-	-	-
Subject peak	.000	.908	-	-	-	-	-	-
Subject excursion	-	-	.000	.833	.002	.732	-	-
Object peak	.000	.892	.000	.621			(.037)	.337
Object excursion	-	-	(.019)	.420	-	-	-	-
Δ object peak/ subject peak	-	-	.000	.622	(.040)	.427	-	-
Δ object excursion/ subject excursion	-	-	.003	.467	-	-	-	-

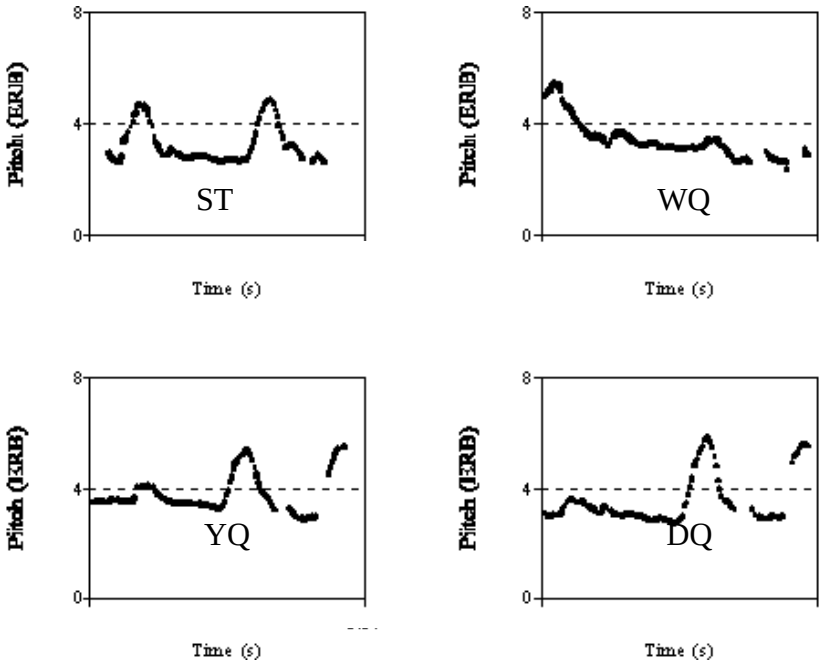


Figure 4.15. Prototypical accent patterns.
Four versions of the core sentence *maRIna wil haar mandoLIne verkopen*: ST, WQ, YQ, DQ. Speaker MC (male).

4.4 Summary of results and conclusions

Sections 4.2.2.2-5 have brought to light a number of *qualitative* differences in accent patterning between statements and questions; in some respects, the three question types were also found to differ between themselves. The principal observations were:

- In statements as well as in questions, the high(-falling) pitch accents H*L and H* dominated; the low(-rising) pitch accents L*H and L* were relatively rare and largely speaker-specific;
- Questions tended to have more high pitch plateaus (H*) than statements;
- Initial boundary tones were predominantly labelled Low (%L), in the statements as well as in the questions, in spite of the questions' higher onset f0;
- In YQ and DQ, pitch of object accents was typically raised relative to the pitch of preceding subject accents ('upstep'), a phenomenon largely absent in ST and WQ; the occurrence of downstep was mainly restricted to ST;
- In WQ, there was a strong tendency for subject and object accents to be auditorily subordinate to the high initial accent on the wh-word.

The *quantitative* evaluation of the accents in the corpus material (§§ 4.3.3.2.1-8) yielded the following results:

- In all three question types, mean subject excursions were significantly reduced relative to those in the statements. Statistically, this parameter had the strongest effect (η^2 =.833).
- In YQ and DQ, mean intervals between object peak and subject peak (Δ peaks) had positive values, reflecting upstepping pitch. The difference with ST was significant. In ST and WQ, values on this parameter were negative, indicating that, between the two peaks, pitch fell.
- In YQ and DQ, mean object peaks had significantly higher f0 values than in the statements.
- In the questions, mean object excursions were larger than mean subject excursions (Δ excursions). However, only the difference between ST and YQ was significant.
- Mean subject peaks did not show significantly lower levels in the questions than in the statements.
- In YQ and DQ, mean object excursions were larger than in the statements, in WQ they were smaller; these differences were not significant, however.
- In WQ, there was a tendency for the peak on the wh-word (wh-peak) to be higher than the subsequent subject and object accents;
- On the whole, there was strong support for the Functional Hypothesis predicting that absence of non-prosodic (i.e. lexical and/or syntactic) question-markers increases high(er) pitch.
- Across the board, the female speakers made larger accent excursions than the male speakers. Differences did not meet the significance level of .01, though.

The next two sections discuss the results against the background of the two research questions underlying the present chapter, viz.

(i) Did speakers raise the nuclear accents in their questions (hypothesis iv) and, more generally, which of the acoustic parameters was most systematically exploited for signalling differences between the four utterance types?

(ii) Did the global regression lines observed in chapter 3 reflect the accentuation patterns typical of each of the four utterance types?

4.5 Discussion

4.5.1 Raised nuclear accent

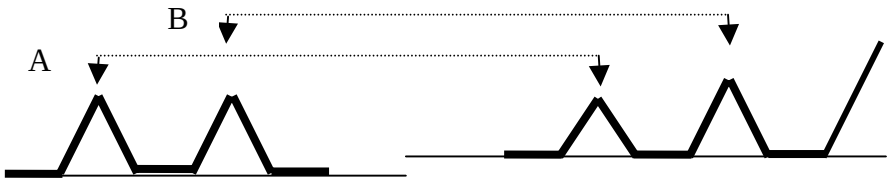
Hypothesis (iv) was tested with the help of the following parameters: WH-PEAK, WH-EXCURSION, SUBJECT PEAK, SUBJECT EXCURSION, OBJECT PEAK, OBJECT EXCURSION, Δ PEAKS and Δ EXCURSIONS. Four of these displayed significant effects at the .01 level, viz. SUBJECT EXCURSION, OBJECT PEAK, Δ PEAKS and Δ EXCURSIONS. Among these, SUBJECT EXCURSION clearly stood out because (i) it had the highest η^2 (.833), reflecting a strong effect, (ii) all three question types differed significantly from the statements, and (iii) the male and female patterns were highly similar, suggesting a certain robustness.

SUBJECT EXCURSION was jointly followed by Δ PEAKS (η^2 =.622) and OBJECT PEAK (η^2 =.621), both of which showed significant differences between, on the one hand, ST and WQ, and on the other, YQ and DQ (in fact, this proves a recurring dichotomy in the four utterance types, see further chapter 6). Along these two parameters, male and female patterns were, again, very similar. Finally, in Δ EXCURSIONS (η^2 =.467) only the difference {ST~YQ} reached significance. That is, of the two Δ parameters, Δ PEAKS differentiated better between the various utterance types than did Δ EXCURSIONS (cf. Figs 4.12 and 4.13). This was probably due to the fact that, in DQ, the subject excursions showed less reduction than in WQ and YQ.

The finding that SUBJECT EXCURSION had the strongest effect would seem to indicate that the reduction of this pitch accent was the most important acoustic property of interrogativity, more important than the (raised) peak f0 of the nuclear object accent, or the difference between subject and object peak. However, at this point it is not obvious whether, overall, the results should be interpreted in a syntagmatic or in a paradigmatic light. On the one hand, looking at the ensemble of the strongest three effects, we might conclude that YQ and DQ require a

(syntagmatic) upstep in f0 between subject and object peak[14]. It could then be argued that the main aim of the reduction of the subject accent in the questions is to enhance the object accent[15]. This would seem to receive support from the results of a perception experiment (Van Heuven & Haan 2002, to appear), which had subjects listen to successively larger stretches of the same declarative question ('gating technique'). Among other things, the stimuli had been varied with respect to the relative excursion sizes of the subject and object accent. Results indicated that the decision as to whether the utterance was a statement or a question was taken only after listeners had heard the object accent and had been able to compare its excursion size with that of the preceding subject accent. In other words, in our experiment the paradigmatic parameter SUBJECT EXCURSION may yet have had a syntagmatic function. This interpretation is in line with observations by Liberman & Pierrehumbert (1984), Rump & Collier (1996) and many others that syntagmatic relations between accents are meaningful (cf. § 4.3.2.1). Obviously, under this view the specific 'meaning' underlying the accent relations observed in our *question* material has yet to be established.

On the other hand, the result on SUBJECT EXCURSION can also be viewed as purely paradigmatic. Speakers may have reduced their subject accents in questions for pragmatic reasons, thereby unwittingly causing the upstep effect. In fact, the raised object peaks in the questions and, consequently, the peak asymmetry reflected in Δ PEAKS may well have arisen from the raised register level, as Figure 4.16 illustrates.



(a) **STATEMENT** (b) **QUESTION**

Figure 4.16. Subject and object accents in (a) statements and (b) yes-no and declarative questions. The thin solid lines indicate register levels, the double arrow A compares the peak levels of the subject accents, B those of the object accents

While the accents in the statements are roughly equal (a), in the questions the excursion on the subject accent is reduced (b). However, because of the raised register level in (b), the *peaks* of the subject accents in (a) and (b) are yet of equal

[14] Obviously, such an account excludes WQ (whose Δ PEAKS and OBJECT PEAK patterned with ST). This is only to be expected, given that the accent pattern in this question type was typically left-asymmetric, rather than right-asymmetric as in YQ and DQ.

[15] Cf. Ladd (1993a), who found that when a nuclear peak is already quite high, listeners perceive it as even more prominent when the f0 of the preceding accent is *reduced*.

height. This may have caused the differences between the subject peaks in the statements and the questions to fall short of significance. By the same token, the object peaks in the questions may have been boosted by the raised register level, their excursion sizes remaining constant. As Table 4.15 showed, the object excursions in the questions (YQ and DQ) were indeed not significantly larger than in the statements. Indeed, the interaction between register level and peak height would seem to make peak height a less reliable parameter when used across statements and questions.

On balance, though we found clear evidence for hypothesis (iv) in that the nuclear accent *peaks* in YQ and DQ were significantly raised relative to those in ST *and* relative to the preceding subject peaks, the mechanism underlying this raised pitch is less clear. On the basis of the present data, it cannot be decided whether the higher nuclear peaks resulted from a globally raised register level, from a locally raised nuclear peak, or from a combination of these two. To establish this, more detailed experiments are required. Nonetheless, taking our cue from the results in Van Heuven & Haan (2002, to appear) who varied relative peak levels but *not* overall register level, we assume that the asymmetric accent patterns in the questions are perceptually relevant. In chapter 6, we explore whether it is possible to attribute these patterns to specific focus structures.

4.5.2 Regression lines vis-à-vis accentuation

In the first part of our investigation of question intonation, we concentrated (i) on *global* acoustic properties (register level, overall trend), and (ii) on acoustic properties of *local* boundaries (utterance onsets, final rises). To establish the global properties, we calculated regression lines and compared their slopes across the four utterance types. These regression lines included the full set of 800 utterances, regardless of possible phonological differences and inclusive of the 17 non-neutral realisations. The purpose of comparing the regression slopes across the four utterance types was to establish whether the global trend of, notably, the lower f0 values acoustically differentiates questions from statements. We assumed that this would be the case (i) when the three question types systematically displayed shallower slopes than the statements without the two categories overlapping, and (ii) when the question slopes adhered to the Functional Hypothesis in that WQ had the steepest downward slope, DQ the shallowest. Crucially, however, question slopes and statement slopes were found to overlap. Besides, the slopes also tended to vary as a function of utterance length and of the presence/absence of a final rise. This made it questionable for global slope to serve as an independent device for distinguishing between the functions declarativity and interrogativity. We speculated that, instead, the global slopes might have been products of the accentuation patterns typical of the four utterance types.

The present chapter focussed on these accentuation patterns. A transcription using ToDI yielded phonological (surface) structures, and statistical analyses indicated whether or not the observed phonetic differences between the utterance types were significant. The results seem to suggest that the regression slopes indeed reflected the accent patterns typical of each of the four utterance types. The circumstance that, in YQ/DQ, the main accent occurred utterance-finally

whereas in WQ it occurred on the utterance-initial syllable no doubt strongly affected the global slopes. The same holds for other differences between the four utterance types such as the presence vs. absence of high plateaus, a small vs. large interval between subject and object peak, accented vs. deaccented subject accents, and the occurrence of 'high H%' vs. 'low H%'.

In the quantitative analyses reported in this chapter, we carefully considered the phonological variation established by the transcription and factored it out whenever necessary. Hence, calculations of means involved ever-varying subsets of datapoints. For instance, from the calculation of the accent minima all accents labelled 'high plateau' were excluded. Or, when calculating, e.g., mean accent excursions we divided the pitch accents into an H*(L) set and an L*(H) set and left the non-prototypical L*(H) outside further consideration. Such selections made the various data sets smaller but at the same time phonologically homogeneous. Besides, the transcription now enables us to compare the slopes of the full set of 800 regression lines with a finer-grained subset from which phonological variation has been removed. This comparison should bring out the remaining differences between the four utterance types which, it must be assumed, are largely *phonetic*. After elimination of the contrastive realisations, the L*(H) accents and all high plateaus we are left with a subset of 483 utterances (150 ST, 127 WQ, 116 YQ and 90 DQ) [16]. Figure 4.17 plots the upper and lower regression slopes of this subset against those of the full set of 800

[16] Questions lacking a pitch accent on subject and/or object were not discarded. Presumably, the absence of a pitch accent had pragmatic reasons. Phonetically as well as conceptually, the difference between no accent and a strongly deaccented accent seemed gradual rather than categorical. That the boundary between presence and absence of accent was often fuzzy was already apparent during the transcription; it was the most frequent cause of disagreement among the transcribers (cf. § 4.2.1).

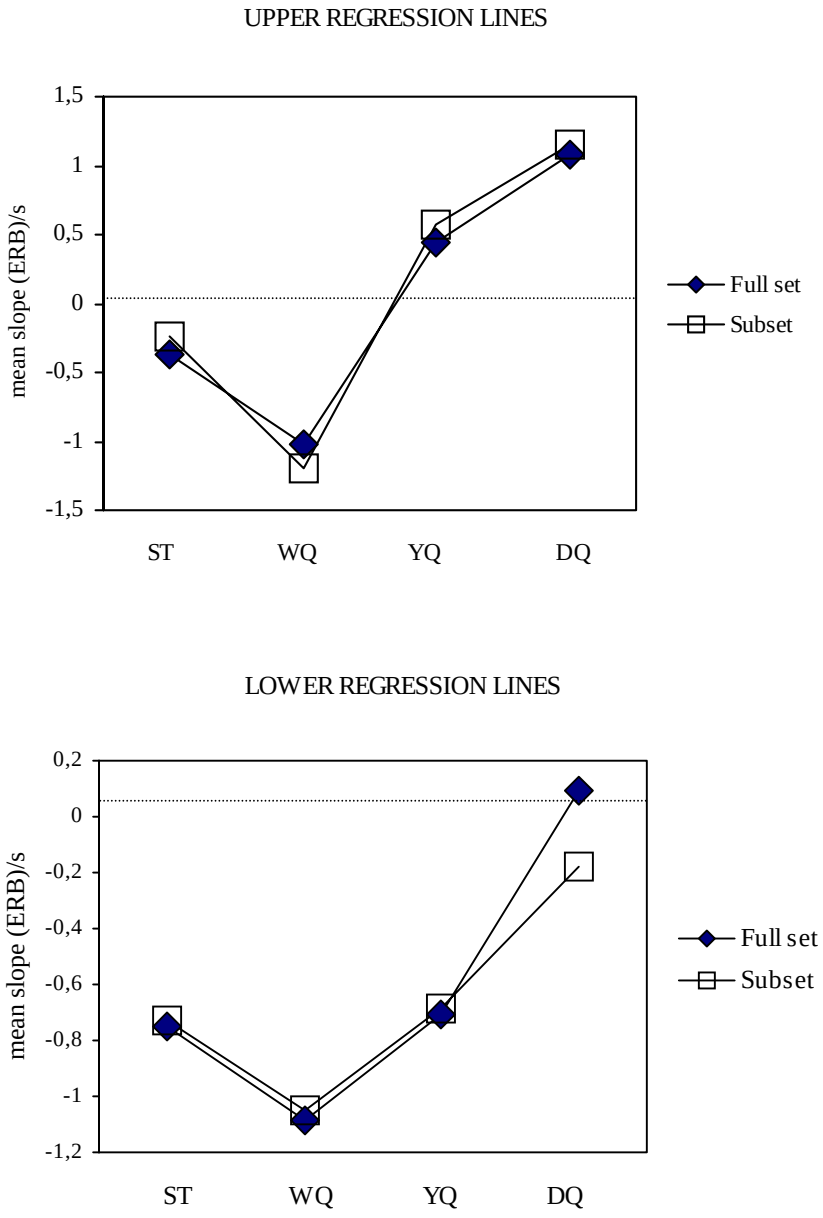


Figure 4.17. Mean slopes (ERB/s) of upper regression lines (upper panel) and lower regression lines (lower

panel) of the phonologically homogeneous subset of 483 utterances, plotted against the full set of 800 utterances and broken down by utterance type. Negative values reflect downward slopes, positive values upward slopes. Means are collapsed over the sexes.

The two graphs show that the pattern of slopes was fairly robust. Although the mean slopes in the full set of utterances contained a certain amount of noise, elimination of this noise did not lead to major shifts in the differences between the four utterance types. In the reduced subset, the upper regression slopes of WQ are still the steepest downward, steeper than those of ST, whereas those of YQ and DQ have, once again, upward courses. As for the lower regression slopes, in YQ and DQ these are less downward than in WQ and ST, in the full as well as in the reduced subset. The greatest difference between the two sets can be found in the lower regression line of DQ, whose slope was clearly more downward in the reduced subset than in the full set (where it was virtually horizontal). This must be due to the fact that the high plateaus were excluded from the reduced set. Notably DQ showed a high incidence of the plateau H* H%, a high-pitched configuration that always occurred utterance-finally. Its presence in the full set no doubt gave the lower regression slope of DQ a more upward slant.

Summarising, the regression slopes in chapter 3 seem to have varied as a function of phonological choices typical of each utterance type (e.g. location of the main accent (WQ), presence/absence and location of high pitch plateaus), together with the phonetic implementation of these choices. Elimination of the phonological part of this variation[17] did not substantially alter the relative differences between the four utterance types. This is taken to mean that, to a considerable extent, these differences were brought about by phonetic implementation, e.g. by reduced H*L vs. expanded H*L ('upstep'), by variation in f0 of the low initial boundary tone %L, or by the phonetic realisation of the postnuclear low. It also means that, by and large, the regression model performed quite well: it succeeded in bringing out global differences between the four utterance types reflecting certain local pitch events. Considering that regression lines are strictly objective and can be calculated semi-automatically[18], the regression model seems well suited for estimating overall trends in large bodies of data. However, for the model to be maximally successful, phonological patterns should not be too heterogeneous.

4.6 General summary

The present chapter compared and contrasted the manifestations of accent in statements and questions. Its primary object was to test hypothesis (iv), according to which questions were expected to display raised nuclear accent peaks. To achieve this, a set of parameters was chosen, viz. peak parameters (f0 values of accent peaks on wh-word, subject and object), excursion parameters (excursion sizes of accents on wh-word, subject and object), and Δ parameters (reflecting differences between

[17] It was evidently not possible to iron out the inherent phonological difference between the three-accent pattern of WQ and the two-accent pattern in the other three utterance types.
[18] I.e., across statements and questions, final rises may have to be identified in advance.

the peaks or excursions of subsequent subject and object accents). Secondarily, with the help of the finer-grained accent data we wanted to evaluate the regression model adopted in chapter 3. Investigations were divided into a qualitative part, in which intonation shapes were identified as a function of utterance type, and a quantitative part in which the four utterance types were compared with respect to a number of numeric parameters. The investigations rested on a transcription of the material, which revealed a certain amount of variation. Although this variation brought some practical problems to the statistical analyses, it also indicated that the range of canonical Dutch (question) intonation is fairly wide.

The qualitative part of the investigation suggested a basic dichotomy within the category 'question', with WQ on one side, YQ and DQ on the other. In WQ, high pitch largely concentrated in the first half of the utterance, that is, in the relatively high-pitched accent on the wh-word and, frequently, in a high plateau following this wh-word; hence, the pattern was left-asymmetric. By contrast, in YQ and DQ high pitch resided predominantly in the second half of the utterance, that is, the pattern was right-asymmetric. This was due to (i) the frequent upstep in pitch between the (commonly reduced) subject accent and the object accent, and (ii) the frequent occurrence of a high plateau between object accent and final boundary tone. A consequence of these different distributions of high pitch across the three question types was that WQ had an overall downward orientation, whereas in YQ and DQ the orientation was mainly upward[19].

Next, the accent data were examined from a quantitative perspective. Assuming that statements represent the default situation, we looked for statistically significant deviations in the questions. One of the main results of this part of the investigation was that it confirmed the existence, in YQ and DQ, of the upstepping mechanism that had earlier been auditorily identified. That is, with respect to these two question types hypothesis (iv) was confirmed[20]. The statistical analyses seemed to suggest that the upstep phenomenon was caused jointly by a reduction of the excursion on the subject accent and the raised object peak of the nuclear accent. That is, the discoursal difference between the utterance types was signalled the most systematically by the parameter SUBJECT EXCURSION, jointly followed by the parameters Δ PEAKS and OBJECT PEAK. Nonetheless, on the basis of the available data it was not possible to establish whether or not the ensemble of reduced subject excursion and raised object peak reflected a syntagmatic mechanism, such that the reduced subject lent greater salience to the subsequent object accent. In an alternative account, the raised object peak may have been wholly or partly due to the globally raised register level. Yet, either way the occurrence of asymmetric accent patterns in the questions was systematic enough to regard it as a prosodic property of interrogativity. Since it is not unlikely for these patterns to have arisen from specific focus structures, this issue is taken up in chapter 6.

[19] It may be repeated that, in the present study discussions of overall trend always concern the question's 'body'; possible effects of final rises are disregarded.

[20] In WQ, subject and object accent were commonly realised on a lower pitch than the (initial) accent on the wh-word, suggesting that they were pragmatically subordinate to the wh-word.

Chapter 5

High question pitch and biological codes

5.1 High(er) question pitch: Q=H

The findings reported in chapters 3 and 4 may be summed up by the formula Q=H, expressing that Questions involve High(er) [1] pitch. Clearly, the strong cross-language tendency for questions to feature high or raised pitch *somewhere in the utterance* (cf. Hermann 1942) also applied in our Dutch question corpus. The findings were in further agreement with Hermann in that they demonstrated that Q=H may take various shapes. Thus, we observed:

(5.1)

(a) *Locally higher pitch,* i.e., the questions showed a greater incidence of intonational forms with intrinsically high pitch (i.e. final rises and high plateaus);
(b) *Globally higher pitch,* i.e., in the questions, register levels were raised and 'upsweep' produced an overall upward orientation;
(c) *Higher accents,* i.e., in the questions, there was a strong tendency for one accent to be raised relative to the other(s). In WQ, this was the question word, in YQ and DQ, the nuclear accent[2]. This resulted in asymmetric accent patterns.

The aim of the present chapter is to gain more insight into the sources of the relatively large quantity of high pitch in the questions. Accordingly, we formulated the following questions:

- What may have caused pitch in the question material to be so consistently *raised* ?
- In the generation of high question pitch, what is the division of labour between the phonology and the phonetics?

As observed earlier (chapter 2, § 2.2.1), the high pitch in the questions may well have been a reflection of a universal biological code, Ohala's Frequency Code. Below, we consider high(er) question pitch largely from the perspective of this code, in an attempt to establish how the latter may have affected the speakers' phonological choices and

[1] As may be recalled, throughout this study, pitch levels in questions are viewed against the backdrop of the supposedly default pitch levels in statements (ST).
[2] WQ stands for wh-question (featuring a question word and inversion), YQ for yes-no question (featuring inversion), and DQ for declarative question (featuring the same lexis and word order as the corresponding statement).

phonetic implementations. Thus, on the basis of the Frequency Code, speakers may have chosen specific phonological categories with a view to increasing the quantity of high(er) pitch. Furthermore, in the phonetic implementation of phonological categories, speakers may have raised pitch in a gradient fashion. Also, we bring in a second universal biological code, Gussenhoven's (1999b, 2002 to appear) Effort Code, whose meanings might, likewise, have played a role in the phenomena listed above.

The chapter is organised as follows. Section 5.2 discusses the Frequency Code and its possible effects on the experimental results. Section 5.3 then explores whether the various instances of Q=H should be analysed as part of the phonology or of the phonetic implementation. First, §§ 5.3.1-2 look at the categorical forms maximising high pitch, i.e. the final rises and the high pitch plateaus. Next, § 5.3.3 concentrates on the raised register level, and § 5.3.4, finally, deals with the accent asymmetry and introduces the Effort Code. Section 5.4 then tries to account for two findings that seemed at variance with our Functional Hypothesis (which was, in part, based on the Frequency Code, see below). These findings triggered a pen-and-paper experiment, which is reported in §§ 5.4.1-5. Section 5.5 concludes the chapter with a general summary.

5.2 The Frequency Code

The consistency of the contrast between the (default) pitch levels in the statements and the higher levels in the questions suggested that this difference is an important prosodic property of interrogativity. This calls to mind Ohala's claim (1983, 1984, 1994) that high question pitch as found in languages all over the world is part of a universal Frequency Code (FC) that must be considered prelinguistic. This code reflects a fundamental opposition between high and low pitch, in mammals as well as in birds, based on the fact that vocalisations are indicative of a vocaliser's size. Accordingly, low pitch denotes 'large', high pitch 'small', with the corollary that pitch also reflects the degree of threat that a vocaliser may pose. Specifically, a large size in a vocaliser is associated with aggression and with competition for the favours of the females, giving the high-low distinction evolutionary overtones. A derived claim is that 'social' messages such as deference, politeness and submission are also signalled by high and/or rising f0, as opposed to authority, confidence and assertiveness, which are conveyed by low and/or falling f0 (Ohala 1994:327). According to Ohala, "the amazing cross-language and cross-cultural similarity of these uses of f0 represent by *themselves* a strong argument for their being innately determined" (1994:331). Asking a question amounts to making an appeal to a listener. This presupposes dependence on the part of the speaker who, after all, relies on the listener's willingness to oblige with an answer. Conversely, the fact that the listener possesses the required information lends him confidence and authority that find their expression in the typically low/falling pitch of his answering statement.

Explaining Q=H as being motivated by the Frequency Code receives support from the fact that the experimental results provided ample evidence for the Functional Hypothesis (FH). As may be recalled, the FH predicted an inverse relationship between the strength of *prosodic* markers of interrogativity and the presence of *non-prosodic* markers of interrogativity (wh-word, inversion). We

expected high(er) pitch to be the most important *prosodic* marker of interrogativity, considering the sub-claim of the Frequency Code that questioning involves uncertainty and dependence on the part of the speaker, expressed by higher pitch. Accordingly, the Functional Hypothesis predicted that, in the three question types under investigation, pitch would become progressively higher in the absence of a question word or inversion. That is, we expected the functionally motivated order WQ<YQ<DQ. As it turned out, this order typically occurred in our data, making it plausible that Dutch question intonation, too, draws on the universal source of the Frequency Code.

At the same time, settling for some universal link between questions and high(er) pitch implies subscribing, at least in part, to what Ladd (1981) has termed the Strong Universalist Hypothesis (SUH)[3]. Ladd's chief objection to the SUH is that the notion of innateness is not compatible with a view that language (and, by implication, intonation) typically avails itself of discrete categories that are fully arbitrary. Adopting the SUH would mean relegating intonation to the domain of paralinguistics. This would be at odds with a view that intonation, like other parts of the grammar, is basically linguistically structured. To support his view that the opposition statement vs. question does not 'naturally' correspond with the opposition low/falling pitch vs. high/rising pitch, Ladd (1996:123) points to Urban North British varieties (UNB), spoken in Birmingham, Liverpool, Glasgow etcetera. Some of the UNB varieties would seem to conflict with the SUH in that they use the same rising-falling nuclear sequence for statements and questions. According to Ladd (1996:125), this is an unambiguous counterexample to the claim that statement intonation is universally falling (and, by implication, that question intonation is universally rising). This, however, does not rule out that UNB questions may yet be prosodically distinct from the statements: it might well be the case that they are realised on a higher overall register level[4], as happens in many languages. Also, the high tone (H) in the statement/question contour L*..H..L% may well be higher in the question version than in the statement. As a matter of fact, when, later on, Ladd returns to UNB intonation (p. 144) he suggests that, at least in Glasgow, the height of this H tone acts as a cue to interrogativity: "[…] the higher it is, the more likely the utterance is to be perceived as a question." As will be seen below, it is precisely in such ways that the Frequency Code may phonetically compensate for phonological form which, for whatever reason, goes *against* the universal trend.

Gussenhoven (1999b, 2002 to appear), on his part, explicitly recognises the force of the Frequency Code which, he notes "we will find it difficult to erase from our repertoire of interpretative strategies". At the same time, he acknowledges that

[3] Part of this Hypothesis as formulated (and refuted) by Ladd (1981) is quoted here: "The SUH makes the following claims: (i) the linguistic functions of intonation are innately specified and/or respond to natural physiological states of the speaker; […….], (iv) Phrasing and sentence-types are signalled primarily by high or rising pitch at the end of incomplete or unresolved phrases or utterances, and low or falling pitch at the end of complete or final ones, [….]. The use of high or rising terminals for "question intonation" signals incompleteness or lack of resolution at the discourse level, with one speaker inviting resolution (i.e. response) from the other speaker." For further discussion of this issue, see Ladd (1996:113 ff).

[4] In the instantiations of a Glasgow statement and question (Ladd 1996:124, figure 4.1) the question is actually spoken on a higher register level; however, the two utterances may have been spoken by different speakers.

languages may feature phonological forms in their questions which must be regarded 'unnatural', since they run counter to the Frequency Code. Such forms may have resulted from language-specific changes, and it seems worth investigating which conditions allow the development of phonological forms that deviate from what the universal FC predicts[5]. Thus, Gussenhoven presents examples of Chickasaw, which has an 'unnatural' reversal of H and L tones (statements end high, questions low), and of Hungarian where both questions and statements end with L. He argues, however, that such cases of 'unnaturalness' can be reconciled with the Frequency Code, because the latter's universal meanings can still be expressed in the *phonetic* space. Although a language basically rests on linguistic structure which lists the phonological forms and their corresponding functions or meanings, the density of these discrete categories is not so high as to take up the entire phonetic space. This offers the speaker room for conveying gradient meanings when phonological categories are given their phonetic implementations.

Gussenhoven's view is in line with the crosslinguistic observations made in chapter 2 (§§ 2.2.2-3). It was noted there that, if questions do not comply with the high(er) pitch requirement, this is usually only partially; elsewhere, Q=H is likely to assert itself. Thus, Chickasaw questions have also been reported to display a raised register level, which can be interpreted as a compensation for their 'unnatural' low endings (Gordon 1999). Moreover, in the Chickasaw *statements*, half of the speakers was heard to realise some slight fall after the 'unnatural' high boundary tone H%. This final fall may have largely or wholly overridden the perceptual effect of the H%. Also, Ladd's observation that UNB statements and questions are phonologically alike yet seem phonetically distinct in fact nicely illustrates Gussenhoven's point. Apparently, within the same phonological category it is possible to convey meaning differences by means of phonetic variation. This variation is consistent with the universal Frequency Code and, if necessary, it may counterbalance 'unnatural' phonological form[6].

5.3 Q=H: Phonetics or phonology?

It seems thus reasonable to assume that the higher pitch in the questions was, to a large extent, motivated by the Frequency Code. Nonetheless, we want to gain a more precise understanding of the contribution of the FC. Therefore, we will look in more detail at the manifestations of Q=H listed in § 5.1, addressing the question of whether these should be seen as part of the phonology or, rather, arise during the phonetic implementation.

5.3.1 Final rises (H%)

[5] As a matter of fact, for some languages Gussenhoven (2002 to appear) offers speculative accounts of the development of falling intonation in questions.

[6] One such type of gradient variation conveying meaning differences is reported by Post (2000). Listeners judged an early timing of the starting point of French final rises as less assured and more surprised than a late timing. Given that an early starting point caused the overall proportion of high/rising pitch to increase, these results were, likewise, in agreement with predictions deriving from the FC.

Of the 583 neutrally realised questions, 87% ended with a rise in f0. In the 200 statements, not a single final rise occurred. There could be little doubt about the categorical nature of this rise: it was either present or absent, and in the transcription this decision never proved problematic[7]. In point of fact, the high boundary tone H% may well represent a grammaticalisation of the Frequency Code (cf. Gussenhoven 1999b:291). Originally, this biological code may have induced speakers to make final high or rising pitch a gradable attribute of a particular set of attitudes. From there, it may have developed into the discrete category H% and become part of the phonology. Synchronically, the choice for this category would no longer seem *directly* motivated by the Frequency Code: despite its non-arbitrary origin it must, by now, be seen as arbitrary.

We may assume, however, that the category H% has broadly retained its original meaning, semantically opposed to the meaning of a final tone that is falling or low. H% has been interpreted as signalling deference, dependence, uncertainty, non-finality, incompleteness, openness, forward-lookingness, continuation dependence, etc. (e.g. Bolinger 1978, Cruttenden 1981, Ohala 1983, Pierrehumbert & Hirschberg 1990, Bartels 1999). Although these meanings are fully compatible with interrogativity, H% would yet not appear an indispensible structural property of questions. Indeed, notably in (English and German) wh-questions, the presence of a final rise is regarded as a marked option. Conversely, H% is not uncommon in *statements* where it may signal continuation, or, at the level of discourse, turn-keeping (Dutch: 't Hart 1998, Caspers 1998, 2000a, Van Donzel 1999; for a discussion of pragmatic aspects of L% vs. H% in English statements and questions, see Bartels 1999). We may speculate that, in WQ and YQ, the main function of the category H% is to make the utterance sound more prototypically questioning, not to signal questionhood *per se*, given that this latter purpose is already served by question word and/or inversion. That is, H% can be seen as a powerful tool for giving expression to the speaker's dependence on the listener's goodwill. However, this dependence may be overruled by some other attitude the speaker may wish to convey along the lines of the Frequency Code (e.g. impatience, anger, irritation, reassurance) and which may require low or falling final pitch. This could cause the speaker to dispense with the final rise although, crucially, he still expects a reply. Only in DQ would H% seem compulsory, given the lack of question word and inversion. In our data, this question type never went without a final rise whereas both WQ and YQ did[8]. Apparently, in DQ the other prosodic features such as raised register or raised nuclear accent peak were not powerful enough to act as a substitute for the final rise.

Earlier on, we regarded the lexico-syntactic markers of interrogativity as factors increasing a question's prototypicality (cf. chapter 1). We could say now that the presence of H% further adds to a question's prototypicality, in YQ but even more so in WQ. On the one hand, in various studies the presence of a final rise in WQ has been associated with attitudes such as more questioning, more polite, friendlier, or more

[7] As noted before, H% could take different shapes. Thus, 'low H%' started from a phonologically low level, whereas 'high H%' took off from a phonologically high plateau; this formal difference is discussed in the next section (§ 5.3.2).

[8] In WQ, the final rise was absent in 36% of the cases, in YQ in 4% (cf. chapter 3, § 3.5.1).

feminine. On the other hand, Scherer, Ladd & Silverman (1984:1352) found that instances of (German) YQ with a final *fall* were judged strongly challenging, and Crystal (1969:4) has pointed out that, without the final rise, British English YQ's sound more serious and abrupt in their implications than when a rise is present. These observations, all of which are in line with the Frequency Code, would seem to support Bolinger's claim (1978b:502) that "the 'more questioning' a question is, [……], the greater the tendency is toward terminal rise".

An analysis approaching the three question types and their intonation in terms of more or less prototypicality would seem to steer clear of the problems that may arise in strictly binary analyses. According to Pierrehumbert & Hirschberg (1990:287), final boundary tones "convey information about relationships among intonational phrases - in particular, about whether the current phrase is to be interpreted with particular respect to a succeeding phrase or not"[9]. They propose that H% is 'forward-looking', in that it directs the listener's attention to an upcoming discourse unit. In yes-no questions, this 'forward reference' of H% is then interpreted as a *cross*-speaker reference. That is, the projected upcoming phrase is to be spoken by the *listener* rather than the questioner. At the same time, Pierrehumbert & Hirschberg are careful to point out that their view of H% is different from proposals that H boundary tones signal "other-directed" utterances, i.e. utterances designed to elicit a response. This is because yes-no questions commonly display H%, whereas wh-questions are typically realised with L%, though both question types must be seen as "other-directed". Yet, when Pierrehumbert & Hirschberg deny WQ 'forward reference' this would seem to imply that WQ does **not** require a response. As said earlier, in our material WQ clearly always meant to elicit a response, no matter whether or not it ended in a final rise. In fact, any account interpreting H% as signalling 'forward-lookingness' (or something similar) in the face of H%'s frequent absence in, particularly, WQ is in danger of running into trouble[10]. Instead, it seems that, notably in WQ, the presence vs. absence of a final rise may signal a range of attitudinal meanings over and above questionhood *per se*, the latter being established by non-prosodic markers. The same holds for YQ, be it to a lesser extent.

It is important to note that H% also showed *phonetic* effects of the Frequency Code. Across the three question types, the size of H% did not differ significantly as a function of question type. Crucially, however, its onset and offset frequencies[11] varied in accordance with the Functional Hypothesis (and, implicitly, with the Frequency Code). As a result, the final rises of WQ, YQ and DQ were typically realised in successively higher portions of the range, as Figure 5.1 shows.

---

[9] Although Pierrehumbert & Hirschberg's interpretation of the meaning of pitch accents and boundary tones is primarily geared to the phonological inventory of American English (as proposed by Pierrehumbert and her colleagues), their compositional approach to intonational meaning and their discussion of discoursal meanings of boundary tones do not necessarily relate to the Pierrehumbert system.

[10] Bartels (1999:219) overcomes this problem by giving the phrasal boundary tones L- and H- (which precede H% or L%) a crucial role in distinguishing between statements, questions and their various (sub)types. This is possible because Bartels, who uses the framework of Pierrehumbert & Beckman, transcribes both H* and H*+L as simple H* (p. 19).

[11] Also termed Postnuclear Low (PL); for a discussion about the terminology, see chapter 4, § 4.3.3.1.

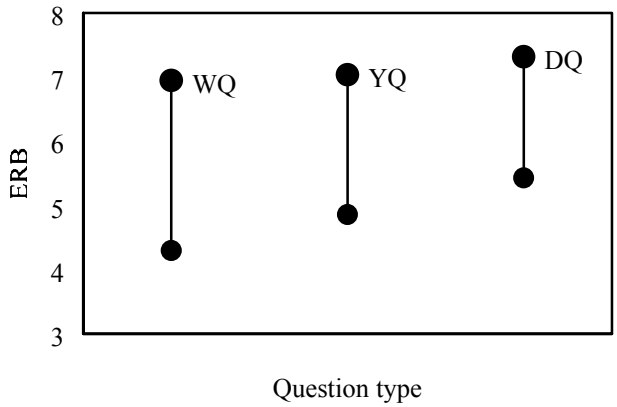


Figure 5.1. Position of final rise within speakers' overall range, broken down by Question type[12].

Clearly, the speakers phonetically differentiated between the three question types, giving the highest pitch to DQ whose question status depends entirely on intonation. Thus, besides using the conventionalised phonological category H% (a fossilised reflection of the Frequency Code), they exploited the remaining phonetic space (i.e. pitch range) to directly express additional meaning deriving from the FC. This fully accords with Gussenhoven's claim that the Frequency Code is a permanent presence in speakers' minds, ready to exert its influence whenever this is expedient. In the case of the final rises in questions its influence would seem twofold, i.e. indirect at the phonological level, as well as direct at the phonetic level.

5.3.2 High plateaus

The transcription results of chapter 4 showed that the questions also featured a relatively large number of high plateaus. These occurred in different positions: between the accent peak on the wh-word and the object peak, between the subject peak and the object peak, and between the object peak and the high final boundary tone. There were also some instances of a low plateau, i.e. between a low starred accent tone and the high final boundary tone.

In ToDI, these plateaus are transcribed as the (spreading) singleton pitch accents H* or L*. However, as pointed out in chapter 2, § 2.1.2, the ToDI system draws on the phonological *surface* level of Gussenhoven's autosegmental description of Dutch intonation; it is not concerned with the *underlying* phonological level (cf. chapter 2, Figure 2.1). At this more abstract level, basic phonological units may undergo modifications such as downstep or tone-linking, the products of which then appear at the phonological surface level. For instance, Gussenhoven's (1988, 1992) Tone Linking

[12] The figure shows the means of all final rises. When occurrences of high H% are eliminated, or when results are broken down by Sex of speaker, the patterns remain virtually the same.

Rule deletes the L tone from a prenuclear H*L pitch accent and links the prenuclear H* to the H* of the nuclear accent (cf. chapter 2, § 2.1.2.2). It seems possible to view all plateaus observed in our question material as linked versions of the basic pitch accents H*L and (relatively infrequent) L*H. Given that the low plateau occurred relatively rarely, this analysis would then imply that the large majority of our contours, unlinked as well as linked, statements as well as questions, ultimately shared the same underlying phonological specification H*L H*L (H%). This would agree with the claim that the pitch accent H*L is "the prototypical, unmarked pitch accent in Dutch, which can be used in almost any situation" (Van Heuven & Kirsner 1999:91, and further references there). A recent transcription of pitch accents in spontaneous Dutch dialogues by Caspers (2000b) confirms this. Of the 1089 pitch accents labelled according to the ToDI system, only 7.5% were labelled 'low' (L*(H)). Of the 92.5% high accents, 70% was H*L, 22.5% H*[13]. In other words, in Dutch the low pitch accent seems a marked option, in spontaneous as well as in read speech. An analysis claiming that, at the *underlying* phonological level, utterances with and without high plateaus consist of the same sequence H*L H*L is not only more parsimonious, it also does justice to the observation that the basic Dutch pitch accent is H*L.

One could think of several reasons why, in the process of speaking, speakers might want to modify the basic phonological units. According to Grabe (1998:93), such adjustments may serve various communicative purposes, i.e., they may affect meaning or focus structure, or they may influence the structure of the discourse. Building on Gussenhoven's modifications and linking rules, Grabe postulates a unified set of categorical phonological adjustment rules which may apply to all pitch accents in an utterance, regardless of their position in the intonation phrase[14] (1998:92). One of her adjustment rules, DELETION, harks back to Gussenhoven's Tone Linking Rule, but is extended to include both prenuclear and nuclear accents. Following Gussenhoven and Grabe but carrying their proposals a step further, we might regard all plateaus in our material (i.e. the forms with the ToDI labels H* or L*) as phonologically adjusted versions of the pitch accents H*L or L*H, that is, as products of a DELETION rule. Below, we illustrate the generation of the various plateau types.

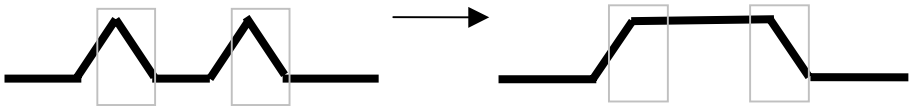
First, we reprint the Tone Linking rule, given in chapter 2, as (5.2a,b). (5.2a) shows a sequence of two basic H*L pitch accents. (5.2b) shows the result of DELETION which has erased the low tone of the first accent and linked the remaining H* directly to the H* of the second accent, creating a high plateau (boxes broadly indicate the CV-part of the accented syllable). According to O'Shaughnessy (1979:143), this f0 pattern "represents a non-contrastive option for the speaker not to mark a boundary within the clause, and may be related to the lack of a contextually-required boundary cue, or to the desire by the speaker to expend less effort (less f0 movement, especially rises, corresponding to less "effort")". In our data, the factor 'effort' seems to have played a role, considering that DELETION occurred more frequently in the short statements than in

[13] Possibly, at the underlying phonological level these occurrences of H* might be analysed as H*L, too (see further below).

[14] Gussenhoven's original proposal distinguished between 'linking rules', which applied to prenuclear accents, and 'modifications', which applied to nuclear accents.

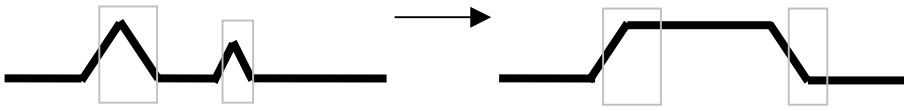
the long ones[15] (30/100 against 11/100, respectively). While a motive for applying the DELETION rule may thus simply be to save effort, the effect is a greater cohesion between the referents of the linked accents (cf. Gussenhoven 1984a). According to Keijsper (1984:36), the meaning of a (Dutch) linked contour ('flat hat') is unlike the meaning of two separate H*L accents, in that the first accent (H*) always announces that another accent is to follow. This, likewise, points to a stronger cohesion between the referents of the accents.

(5.2a) (5.2b)



Similarly, we assume that DELETION accounts for the (b) forms given below. In (5.3a,b), the basic prenuclear accent (H*L) is linked to a downstepped nuclear accent (!H*L), with downstep being carried over onto the resulting 'flat hat' (cf. Gussenhoven & Rietveld 1992). This downstep causes the high level pitch to fall earlier than in the case of a non-downstepped H*, which is why it suggests a different type of fall[16]. In the same way, (5.4b) is seen as the result of DELETION in a sequence of a basic prenuclear accent and an *up*stepped nuclear accent (5.4a).

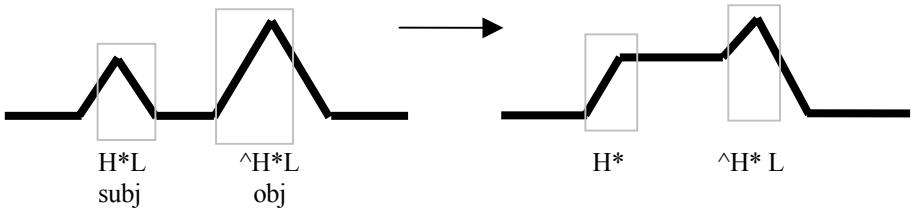
(5.3a) (5.3b)



(5.4a) (5.4b)

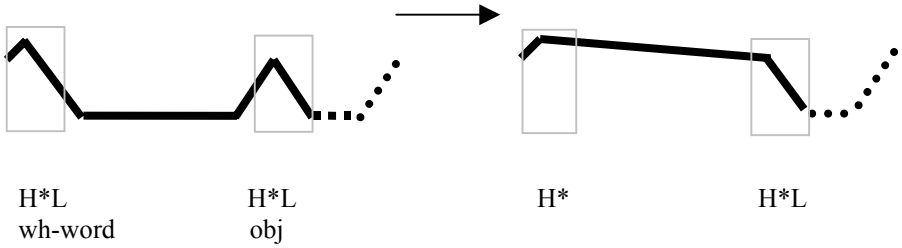
[15] In the short statements, there were only two unaccented syllables between subject and object, whereas in the long ones there were as many as five.

[16] In a perception experiment, Caspers (1999) found that the early fall in %H !H*L differs from the later fall in %H H*L, in that it is not suitable for highlighting new information.

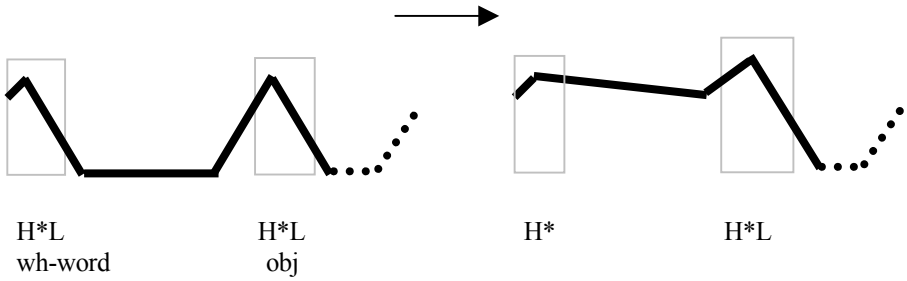


Moreover, DELETION is taken to have applied between the initial peaks on the wh-words and the object peaks, as shown in the variants (5.5a,b) and (5.6a,b). Obviously, in such cases, there was no accent on the subject. In (5.5b), the transition between the (higher) first accent peak and the (lower) second accent peak is smooth. In (5.6b), the first and second accent peaks are more or less equal, causing the transition to slightly 'sag'. Despite these phonetic differences, the phonological strings underlying (5.5b) and (5.6b) are identical.

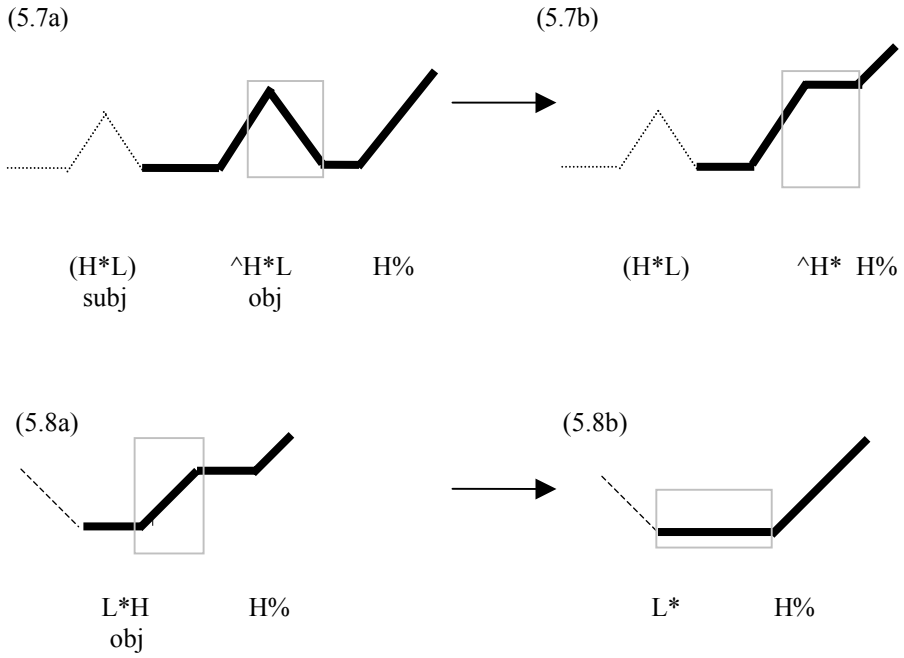
(5.5a) (5.5b)



(5.6a) (5.6b)



The last type of plateau occurred after the object accent and amounted to a direct linkage of the starred accentual tone to the (high) boundary tone (cf. 5.7a,b and 5.8a,b below).



Analysing the (b)-forms as resulting from DELETION implies that the scope of this phonological adjustment rule has to be extended to include boundary tones. One may object that such a move ignores the fact that accent tones and boundary tones represent functionally different classes. Yet, justification for treating these two alike may be found in Gósy & Terken's (1994) perception study of Hungarian question intonation. According to Ladd's (1983a) phonological analysis of Hungarian, the last stressed syllable before the penultimate syllable requires a low tone. With this in mind, Gósy & Terken manipulated a naturally-produced Hungarian sentence. One of their variables involved substituting the obligatory low tone before the penultimate syllable with a rise-plus-fall, on the expectation that this high accent peak would exclude a question interpretation. Although this proved correct in many of the cases, listeners nevertheless also assigned 'question' labels to utterances featuring the high peak. Crucially, however, for such utterances to be interpreted as questions the additional accent peak had to be lower than the subsequent high peak of the HL boundary tone. Conversely, raising the additional high accent peak relative to the boundary peak decreased the number of question responses. Apparently, listeners balanced the f0 of these peaks against each other. Gósy & Terken (1994:279) conclude that such comparisons may involve functionally different peaks, i.e. accent highs as well as boundary highs. In fact, our own production material would seem to support such a view. As observed earlier (chapter 3, § 3.6.1), speakers tended to give the offset of the high boundary tone the highest f0 of the entire question, that is, its level exceeded the functionally different peak(s) of the

preceding accent(s). Apparently, pitch relations between consecutive peaks affect the interpretation of a message, no matter the function of these peaks. Hence, it seems reasonable to assume that functionally different tones, i.e. accent tones and boundary tones, can be linked up to form plateaus.

From a formal point of view, including all pitch plateaus into a single DELETION analysis is no doubt attractive. On the assumption that, underlyingly, pitch accents are sequences of two tones the latter of which may be deleted, the phonological adjustment rule applies across the board, i.e. to all accents regardless of their position or function in the utterance (cf. Grabe 1998). This comprehensive DELETION rule then accounts for the various pitch plateaus observed in the material. Nonetheless, the question arises as to what should motivate such a rule. Is it an 'empty' phonological rule, comparable with, say, the optional segmental rule which deletes /τ/ from Dutch /jOΞτπ↓Σ↔σ/ ('pimples')? Surely, such cases of 'free' variation may convey certain stilistic differences, but there seems no reason to think that this is true of the DELETION rule under discussion here. In fact, a comprehensive DELETION analysis would be more convincing if it could be shown that the deletion of trailing tones represents some constant meaning. As a matter of fact, this was investigated for the forms exemplified in 5.7 and 5.8 above. In a perception experiment, Haan, Heijmans, Rietveld & Gussenhoven (2000) tested the hypothesis that DELETION would give rise to identical pragmatic effects across H*L and L*H pitch accents. This experiment is briefly reported in the next section.

5.3.2.1 DELETION and 'meaning'.

As will be recalled, in our YQ and DQ question data the sequence (H*L) H*L H% alternated with (H*L) H* H%, i.e. between the nuclear accent and the high final boundary tone, a high plateau could occur (cf. chapter 4). This alternation, illustrated in 5.7a,b, was observed in eight out of the ten speakers, often in renderings of the same utterance in the same condition by the same speaker. The same held for the (low) plateau variant of L*H (5.8b), be it that, across the corpus, the occurrence of the low pitch accent was rather limited. The idea underlying the perception experiment was that if DELETION[17] expressed some invariant shade of meaning, its effect on the H*L and the L*H pitch accent ought to be the same, that is, listeners' judgements should not differ as a function of pitch accent.

A female speaker produced the four relevant contours (i.e. H*L H%, H* H%, L*H H% and L* H%) on four Dutch prepositional phrases. Since each phrase could begin with a high or a low onset, this resulted in 4x4x2=32 stimuli. Listeners indicated on a 10-point scale how appropriate the contours sounded, assuming they needed to express a Question, Surprise, a Reminder, a request for Agreement, a Challenge or a Suggestion. The results produced interactions between the two pitch accents H*(L) and L*(H), suggesting that DELETION does not represent an independent meaningful

[17] Haan *et al.* (2000) used the term 'phonologised undershoot' rather than DELETION. This term expressed that deletion of the accentual trailing tone before H% may have developed out of an intonational 'short-cut' which, originally, was motivated by lack of segmental space but became generalised and developed into a phonological category (cf. Gussenhoven 1999b).

element. Instead, another pattern emerged. It turned out that the pragmatic scales Question, Surprise, Suggestion and request for Agreement received the highest scores when the contours had a maximal quantity of intrinsically high pitch (i.e. %H H* H%); crucially, this did not hold for Reminder and Challenge. For instance, on the Question scale the maximally high contour got the highest mean score of 8.21, whereas the maximally low contour (%L L* H%) elicited the substantially lower mean score of 5.77. What the scales Question, Surprise, Suggestion and request for Agreement have in common is that the speaker solicits the hearer's cooperation and involvement. In contrast, the scales Reminder and Challenge both reflect that the speaker is in a position of authority. The finding that the former four scales elicited better scores for appropriateness when the stimuli had high pitch levels, whereas the latter two scales did not, suggested a connection with the Frequency Code. Clearly, as a corollary of DELETION the experimental contours had displayed substantial variation in the overall amount of (intrinsically) high pitch. In retrospect, it seemed likely for the subjects to have based their scores on this overall quantity of high pitch, rather than on the underlying morphological categories.

To verify this, we computed a weighted overall mean f0 for each of the four experimental contours, based on the mean f0 values for the onset stretch, the accented syllable and the two subsequent unaccented syllables. Between these overall means and the listeners' scores, positive correlations were found, that is, higher average pitch corresponded with higher scores. Crucially, however, this was only the case in the scales Question, Surprise, request for Agreement and Suggestion, i.e. the scales expressing a speaker's dependence on a listener's cooperation and involvement. By contrast, the scales reflecting a speaker's authority and assertiveness (i.e. Reminder and Challenge) showed negative correlations, that is, the higher the average pitch, the less appropriate the contour was judged.

The results of this experiment suggested (i) that it is unlikely for DELETION in (5.7) and (5.8) to constitute an intonational morpheme with a well-defined meaning of its own, and (ii) that listeners may base pragmatic judgements of (contextless) intonational stimuli on a non-linguistic code such as the Frequency Code, rather than on linguistic structure. The latter finding would seem to support a view that it is possible for intonation to be substantially informed by the Frequency Code, not only in the phonetic implementation but also in the choice of phonological form. Crucially, the four experimental contours differed only in phonological structure; the phonetic implementation was constant. Nonetheless, the listeners consistently associated the intrinsic pitch levels of the phonological categories with certain attitudes, in ways that fully agreed with the universal FC. At the same time, however, it was also evident that DELETION could not be seen as a phonologisation of the FC, given its opposite effects on overall pitch as a function of pitch accent. For while deletion of L from H*L before H% brought about an increase of the overall quantity of *high* pitch, deletion of H from L*H before H% led to a longer stretch of *low* pitch.

5.3.2.2 High plateaus: Concluding remarks

Given that listeners' judgements of higher-pitched vs. lower-pitched contours clearly corresponded with meanings deriving from the Frequency Code, it seems reasonable to

relate this to our production results. In chapter 4, we reported that 20% of the nuclear tone patterns in the questions using the high pitch accent consisted of H* H% rather than canonical H*L H% (Table 4.11). The incidence of H* H% increased in the order WQ<YQ<DQ, i.e., it occurred most frequently in the question type typically depending on intonation. Likewise, in the (infrequent) nuclear tone patterns using the low pitch accent the sequence L*H H%, which maximised high pitch, displayed the same order: WQ<YQ<DQ. Conversely, with the maximally low sequence L* H% the order was inverted: DQ<YQ<WQ. Chi square tests showed that associations between question type and contour type (maximalised vs. minimalised high pitch) were significant, overall as well as pairwise. What these systematic orders seem to suggest is that a biological code such as the FC may also trigger *categorical* choices. For speakers expressing meanings which the Frequency Code associates with high(er) pitch, categorical deletion of a low tone would then constitute an additional means of raising overall pitch. That this device was used the most frequently in DQ is only to be expected, when we assume that in this question type the dependency implied by questioning is being given maximal expression. However, since we had to conclude that an across-the-board DELETION rule cannot be motivated by the Frequency Code, deletion of low tones as prompted by the FC must be seen as occurring on an *ad hoc* basis.

In sum, it seems clear enough that the formal transparency of a wide-scope DELETION analysis does not correspond with an equally transparent dimension of meaning. The results of the perception experiment showed that DELETION cannot be regarded as an intonational morpheme with a constant meaning which reflects the Frequency Code. Indeed, (5.2b) and (5.3b) pose a further problem for an account claiming that, across the board, DELETION is motivated by the Frequency Code. For although the (b) forms were higher-pitched than the (a) versions, they yet occurred in 20% of the *statements*. However, an important overall conclusion seems warranted: it seems possible for phonological structure to be affected by the biological Frequency Code, such that speakers may give preference to certain categorical forms over others. Options would seem to include more than just the choice between the boundary tones L% and H%, i.e., the FC may also prompt speakers to categorically delete underlying L accent tones in order to increase the surface amount of high pitch in an utterance

5.3.3 Register level

In chapter 3, utterance onset f0, register level (defined as minimum f0) and f0 of the postnuclear low were treated as separate parameters. However, considering that the mean values of each of these were consistently higher in the questions than in the statements[18], it seems reasonable to assume that these higher values in fact reflected an overall raised register level. Therefore, we now collapse them into the single parameter 'register level' which, however, should be kept distinct from the parameter 'register span' since both can be varied independently (cf. Ladd 1996:261). Nor should 'register

[18] Recall that the onsets, which were *phonetically* higher in the questions than in the statements, were only labelled *phonologically* high (i.e. %H or %HL) in 11% of the cases. Conversely, the label %HL was also assigned to 4% of the statements (cf. chapter 4, Table 4.4). Crucially, however, in the questions f0 of the phonologically *low* boundary tones (%L) was higher than in the statements.

level' be confused with the global trend of the lower regression line; in chapter 4, the latter was analysed as a by-product of phonologically specified tonal targets, not as a parameter in its own right.

It is not immediately obvious, however, how a model of intonation should handle raised register level. Is there a discrete phonological category [+raised register], or is register level raised in a gradient fashion? In some descriptions of intonation, register level has been analysed as part of the phonological system. For instance, in their analysis of the tone language Hausa, Inkelas & Leben (1990) model a four-way contrastive distinction between high and low tones in questions by introducing a separate register tier. This makes it possible for H and L tones to be realised on a high vs. a low register. Sosa (1991:120), analysing the intonation of (Latin American) Spanish questions, advocates a phonological analysis of raised register which involves adding a high initial boundary tone to questions. This %H raises the f0 of the first syllable, making it higher than in a corresponding statement. Next, the f0 values of all subsequent tones in the question are generated on the basis of this high initial pitch level. Yet another way of accounting for register shift in the phonology is tentatively proffered by Ladd (1998) to model the observation that the f0 in Danish questions remains well above the bottom of the speaker's range. Ladd suggests that it is possible to interpret this as a binary choice between contours that are [+bottom f0] and [–bottom f0].

A phonological approach might, arguably, imply the joint assumptions that (i) a raised register is indispensible to questions, (ii) that non-questions do not display register shifts, and (iii) that register shifts only *up*wards. These assumptions seem somewhat questionable, though. Some evidence against (i) is provided by a perception experiment where listeners had to decide whether an utterance was a statement or a declarative question (Van Heuven & Haan 2001). Here, listeners overwhelmingly relied on the presence vs. absence of a final rise. It was only when they were prevented from hearing the final part of the utterance that they based their judgements on the other variables 'accent size' and 'declining/inclining global trend'. Since the stimuli did not include raised registers, it would appear reasonable to assume that questions in no way *require* register to be raised. According to Studdert-Kennedy & Hadding (1973:295), there is no upward register shift in questions to which the speaker thinks he knows the answer. Earlier, we argued that such questions ('confirmation questions') could be regarded less prototypical than 'information questions'. It would appear that, in fact, the chief effect of a raised register level is that it causes a question to sound *more prototypical*, along the lines of the Frequency Code. As regards assumption (ii), attitudes or emotions such as 'surprise' and 'fear' are, similarly, realised on raised register levels (Van Bezooijen 1984, Mozziconacci 1998). Moreover, as Gårding (1982:133) points out, overall upward (or downward) shifts in register level between consecutive (Swedish) utterances may express meaningful relations between these utterances. In other words, an upward shift may serve more purposes than expressing the opposition statement-question. As for assumption (iii), it is equally possible for pitch to display a *downward* shift of register, e.g. in utterances spoken in a confidential or conspirational manner, or in 'asides' or parenthetical remarks. This means that, in a model claiming that register level results from a categorical choice, at least three register levels have to be recognised: a default level, and variants which are raised or lowered

relative to this default level.

Altogether, there would not seem much reason for designating a default register level and a raised register level as categorical choices corresponding to the discoursal categories 'statement' and 'question', respectively. It seems more plausible to regard upward/downward modification of the default level as an overall device for expressing or supporting a fairly wide range of gradient non-linguistic meanings (cf. Crystal 1969). What is important is that the broad correlations between register level variation and the types of use this variation is put to suggest a connection with the Frequency Code. Thus, high(er) pitch level goes together with questioning, surprise, fear, whereas overall low(er) pitch is more typical of assertion or imperativity (e.g. Van Heuven & Kirsner 1999).

5.3.4 Asymmetric accent patterns

In the statements (ST) as well as in the wh-questions (WQ), accent peaks of subject and object were found to be more or less equal, that is, neither peak values nor accent excursions significantly differed from one another. By contrast, in the yes-no questions (YQ) and declarative questions (DQ), pitch between subject and object accent frequently stepped up, giving rise to asymmetric accent patterns. This asymmetry was generally caused by a reduction of the excursion of the subject accent, together with a higher peak on the object accent; hence the accent pattern was *right*-asymmetric. Statistically, this pattern differed significantly from the patterns in ST and WQ (cf. chapter 4, Table 4.16). Conversely, WQ often featured a *left*-asymmetric pattern because of a sharp fall in pitch between the relatively high-pitched initial wh-word and the stretch comprising the relatively low-pitched subject and object. Bolinger (1972a:139) approaches unequal-peak relationships by drawing tangents from one accent peak to the next, as illustrated in Figure 5.9. Such tangents, which ignore the f0 of intervening *un*accented syllables, have an upward or a downward direction comparable with overall rising or falling contours.

Figure 5.9. Asymmetric peak relations.
Upstepping accent peaks display an overall rising tangent (a), downstepping accent peaks an overall falling tangent (b). Adapted from Bolinger (1972a:141).

According to Bolinger, in American English a rising tangent has broadly the same pragmatic use as a contour in which all syllables, accented as well as unaccented, are continuously rising (although the two may slightly differ in their appropriateness to situations). The same is true of a falling tangent vis-à-vis an overall falling

sequence of accented and unaccented syllables. Bolinger goes on to claim that a rising tangent represents hearer-orientation and makes an appeal to a listener, whereas a falling tangent represents speaker-orientation and expresses that things are settled (1972a:149).

Clearly, these claims fully accord with predictions made on the basis of the Frequency Code and one would therefore expect questions to display rising tangents[19]. However, while this is true of YQ and DQ, the pattern in WQ constitutes a counterexample. Tangents drawn between the initial accent on the wh-word and the stretch containing the subject and object accent were frequently *falling*, as in 5.9b. This means that the asymmetric accent patterns in the questions cannot be explained on the basis of the Frequency Code. For an account that generalises across all three question types, we have to look elsewhere.

Obviously, whether tangents in our experimental questions were rising or falling followed directly from the position of the largest accent (cf. Eady & Cooper 1986). While in WQ this accent occurred utterance-initially, in YQ and DQ it was utterance-final. This positional difference accounts only partially for the observed asymmetries, though. The crucial preliminary question is why the accents in the questions should at all have turned out comparatively small or large, relative to the 'neutral' accents in the statements. Given that higher peaks correspond with greater perceived prominence (e.g. Rietveld & Gussenhoven 1985, Gussenhoven & Rietveld 1988), accent patterns with equal vs. unequal peaks are bound to receive different interpretations. Indeed, variation in relative accent size is possibly closely bound up with another universal biological code, Gussenhoven's Effort Code. As this code appears relevant for our understanding of the asymmetric accents patterns, the next section considers it in more detail.

5.3.4.1 The Effort Code

Alongside the Frequency Code, Gussenhoven (1999b, 2002 to appear) identifies a second innate mechanism bearing upon intonation, the Effort Code. Basically, this code reflects that the size of a pitch obtrusion is a function of the degree of effort put into it by the speaker and, by implication, of the degree of explicitness the speaker intends to convey. Accordingly, a listener will associate larger pitch movements (and, by implication, higher peaks) with greater explicitness (or: insistence, emphasis, significance). As it is, the Effort Code would seem to go some way towards explaining the accent asymmetry in our question material. The raised object peaks in YQ and DQ may have communicated that the corresponding words (or constituents) were of relatively great significance, whereas the systematic reduction of the subject accents may have expressed that the corresponding words (or constituents) were relatively insignificant. Or, put differently, in the patterns comprising a small subject accent followed by a large object accent, the latter may have carried more weight at the expense of the former. By the same token, the frequently occurring WQ pattern (a high accent peak on the initial wh-word followed by considerably lower subject and object

[19] In the Frequency Code, rising pitch shares a category with high pitch, and falling pitch with low pitch (cf. Ohala 1983).

peaks) may have conveyed that the wh-phrase was comparatively significant. In fact, as the *Anna/Manny* and *Amanda/Malta* experiments have shown (cf. chapter 4, § 4.3.2.1), such asymmetric accent patterns typically reflect *non-neutral focus*. Ultimately the realisation of focus, whether neutral or non-neutral, may be seen as a grammaticalised reflection of the Effort Code, in that the (relative) scaling of the accent mirrors the (relative) communicative weight of the corresponding constituent.

Seen in this light, the observed variability of accent patterning as a function of the opposition statement-question may have reflected differences in focus structure. However, to investigate this we need a better understanding of focus *in questions*, an issue that would not seem to have attracted much attention so far. Therefore, we address it in a separate chapter (chapter 6) which investigates (i) whether focus in questions is different from focus in statements, (ii) whether this difference accounts for the observed variation in the accent patterns, and (iii) whether this variation in accent patterns may, ultimately, derive from the biological Effort Code.

5.3.5 Biological codes: Discussion

Although both Ohala's Frequency Code and Gussenhoven's Effort Code must be thought of as having non-linguistic origins (in that they indicate the vocaliser's physical size, or relate size of pitch movement to physical effort), they may have made their way into the linguistic system and become phonologised. That is, they may have produced distinct phonological categories that have become part of languages' inventories. Diachronically, the boundary between the non-linguistic/paralinguistic system and the linguistic system is likely to have been repeatedly crossed by elements from the former becoming part of the latter. A prime example of phonologisation of the Frequency Code would be the final rise (Gussenhoven 1999b), the chief prosodic attribute of Dutch questionhood, in production as well as in perception (for the latter, see Gooskens & Van Heuven 1995, Van Heuven & Haan 2001). It is less obvious whether phonologisation has also applied to a phenomenon such as emphasis. Assuming that the larger excursion of an emphatic accent reflects the operation of the Effort Code and signals greater explicitness, is there a categorical difference between a 'normal' and an 'emphatic' accent? Results of a categorical perception experiment by Ladd & Morton (1997) have suggested that if the distinction is not truly categorically *perceived*, it would yet seem to be categorically *interpreted.* That is, the listeners showed themselves "predisposed to interpret accents or utterances as belonging to one of either category, "normal" or "emphatic" " (p.339). As for the phonetics, it cannot be doubted that these offer ample evidence of the operation of both the FC and EC. It thus appears that, throughout the intonation grammar, these two biological codes act as strong (non-linguistic) undercurrents: they may underlie phonological categories, govern phonological choices, affect the phonetic implementation and guide the perception. That is, iconicity is less remote from intonation than it allegedly is from other linguistic structures.

The vocal aspect over which Frequency Code and Effort Code have scope is what is usually loosely referred to as 'pitch range'. This may be broken down into (i) the global raising/lowering of register *level*, and (ii) the local increase/decrease of pitch movements resulting in widening/narrowing of register *span*. As a matter of fact, variation in pitch range poses knotty problems for intonation theory, especially for AM

theory which recognises only two abstract tone levels, H and L. In 1980, Pierrehumbert's two-tone description of English intonation provided a strong impetus to the development of the AM framework. A major advantage of a two-level model over the then current four-level model was that, in the latter, it was hard to tell whether a certain f0 value derived from a tonal specification or from a choice of pitch range. This blurred the line between these different sources of variation whose individual contributions to a given f0 value should, instead, be investigated separately. At the same time it was evident that, for a *two*-tone model to work, it needed to model variation in pitch range. Therefore, Pierrehumbert posited a set of context-sensitive implementation rules which automatically alter the phonetic values of H and L as a function of their tonal context. For instance, an H following an H+L accent is lowered (downstep), and a high boundary tone following a H phrase tone is raised (upstep). In this system, variation in pitch range chiefly arises from the choice of pitch accent, together with the automatic implementation rules producing 'allotones'[20].

Clearly, this type of phonetic implementation does not carry any meaning, it simply 'happens', in a mechanical way. This makes it unlike the type of implementation in Gussenhoven's (1999b, 2002 to appear) proposal, which can be employed as a channel for gradient universal meanings (see § 5.2 above). Gussenhoven gives several examples of local pitch range variation expressing this type of meaning. For instance, in the southern Dutch (Limburgian) dialect of Roermond, statements and wh-questions both end in the same falling-rising phonological sequence H*LH. Nonetheless, speakers signal the functional difference between the two utterance types by giving this sequence different phonetic implementations. Thus, in WQ, the rise runs from mid-range to high-range, whereas in ST it stretches between low-range and mid-range. Evidently, this Roermond variation adheres to the Frequency Code, that is, the higher pitch goes to the question, not to the statement. In such ways, speakers may exploit the phonetic implementation to give expression to universal meanings, over and above the intrinsic 'meaning' carried by the phonological category. Following Gussenhoven's proposal for integrating the universalist perspective into intonation theory, we could say that the Frequency Code, broadly, affects register *level*, whereas the Effort Code mainly affects register *span* (although, admittedly, it is not always easy to draw the line, cf. Ladd 1996).

It seems obvious that our production material showed the effects of at least one universal biological code, Ohala's Frequency Code. First, the questions featured more *phonological* forms of intrinsically high pitch than did the statements, viz. final boundary tones (H%) and high plateaus (resulting from the deletion of L tones). Second, the *phonetic implementation* caused phonologically low tones to have higher pitch in the questions than in the statements. Thus, while the low initial boundary tone (%L) was predominant both in the statements (98%) and the questions (89%), its mean pitch was higher in the questions (4.7 vs. 5.1 ERB, respectively[21]). The same was true of the postnuclear lows; here the means were 3.8 ERB in the statements, against 4.5 ERB in

[20] Pierrehumbert's (1980) system also allowed for 'external' sources of pitch range variation, though. Thus, a new choice of pitch range for expressive use was possible at each new pitch accent, depending on the stress subordination in the phrase and/or a speaker's wish to highlight particular information (p.7).

[21] Individual means were: ST 4.7, WQ 5.6, YQ 5.1 and DQ 4.8 ERB.

the questions[22]. Third, the questions had higher minima than the statements (4.2 vs. 3.7 ERB, respectively[23]). In brief, the association between questions and higher pitch was highly consistent, in the phonology as well as in the phonetics. Also, there was a strong tendency for pitch to be highest in the question type that relied most on intonation (DQ), which is precisely what the FC would predict. Finally, a separate perception experiment, including two unlinked and two linked nuclear contours, strongly suggested that listeners' judgements of the stimuli were guided by the FC.

The results seem to warrant the conclusion that the universal Frequency Code operates at two levels: (i) it may lead speakers to prefer intrinsically high phonological categories to categories that are intrinsically low, or *vice versa*, and (ii) it may lead speakers to employ the phonetic space of phonological categories to give additional expression to certain gradient meanings (cf. Gussenhoven 1999b, 2002 to appear). However, for one of the Q=H phenomena listed in § 5.1 the Frequency Code could not account, viz. the accent asymmetries observed in the questions. Although it seems likely that these different patterns rested on different focus structures and derived, ultimately, from a second biological code, the Effort Code, we first have to gain more understanding of possible differences between focus in statements and in questions. This means that the discussion of whether question intonation involves not just the Frequency Code but also the Effort Code, has to be postponed till chapter 6.

However, before concluding these sections on the influence of universal codes we have to account for what may strike us as counterexamples to the FC. Earlier, we reported that the declarative question (DQ) in the experimental corpus seemed to fail the Functional Hypothesis and, implicitly, the Frequency Code on two counts. In order to decide whether these observations could be attributed to functional differences between DQ and YQ, we conducted a pen-and-paper experiment. This will be further introduced and reported in § 5.4 and subsections.

5.4 Information vs. confirmation: Introduction

As stated earlier, the results reported in the chapters 3 and 4 generally provided strong support for the Functional Hypothesis and, implicitly, for the Frequency Code. Thus, the differential use of pitch in the three question types appeared to be functionally motivated, such that the quantity of high question pitch increased as questionhood depended more strongly, or entirely, on intonation. Consequently, in the set of question types, Q=H typically displayed the order WQ<YQ<DQ. Nevertheless, we noted two exceptions to this overall rule. In chapter 3, it was seen that the onset levels of DQ, though higher than those in the statements, were not significantly different from the latter; this held for each of the ten speakers (§ 3.5.2.1). In chapter 4, we observed that the subject accents in DQ were not reduced to quite the same extent as in YQ, despite the fact that, otherwise, these two question types were very similar. Both observations might be interpreted as 'deviations', in DQ, from the Frequency Code. In addition, while YQ featured 32 deaccentuations of the subject, in DQ there were only six. Overall, in DQ the

[22] Individual means were: ST 3.8, WQ 4.1, YQ 4.5 and DQ 4.9 ERB.
[23] Individual means were: ST 3.7, WQ 4.1, YQ 4.3, and DQ 4.3 ERB.

subject constituent seemed to carry somewhat more discoursal weight than in YQ[24]. In effect, the initial portion of DQ resembled that of ST, be it with an upward tilt. Below, Figure 5.10 gives a schematic representation of onsets and subject accents in the two utterance types; the upper contour represents DQ, the lower contour ST.

Figure 5.10

That the f0 of phonologically low onsets (%L) was higher in questions than in statements was regarded as an early and therefore effective speaker device for signalling an utterance's question status. Likewise, a reduced subject accent was viewed as an important (early) prosodic property of interrogativity. That these features appeared somewhat less prominent in DQ raised the question of whether this utterance type might (be intended to) sound less 'questioning' than YQ. Although most authors put DQ on a par with YQ (Van Alphen 1914:91; Den Hertog & Hulshof 1972:142; Haeseryn *et al.* 1997:1428), Droste (1972) is of the opinion that the former seeks confirmation rather than information. This would make DQ functionally similar to a 'hè'-question, a statement to which the high-ending tag *hè?* ('right?') is appended, turning the statement into a question for confirmation (cf. Droste 1972, Kirsner & Van Heuven 1996). One might speculate that DQ is used to express a broadly equivalent function, with the role of the lexical tag *hè?* being taken over by intonation. This might then account for the slightly hybrid prosodic shape of DQ: whereas its first half resembles a statement, its second half is steeply rising, and this strong overall rise might have a similar communicative function as utterance-final *hè?*.

In order to settle this issue and to determine whether the prosodic differences between YQ and DQ in fact corresponded with a functional difference (information question vs. confirmation question), we carried out a pen-and-paper experiment. Design, materials and results are laid out in the sections below.

5.4.1 Design and materials

Design

The notions 'information question' and 'confirmation question' were assumed to occupy opposite poles on a continuum of predictability of the response. That is, the

[24] That subject accents were slightly less deaccented in DQ than in YQ only 'counts' as a deviation from the Frequency Code when deaccentuation of the subject is interpreted as causing the object accent to become more salient. Still, the mean upstep between subject and object accent was larger in DQ than in YQ (although the difference was not significant), which was in perfect accord with the Frequency Code.

former was operationalised as a question the answer to which was thought to be maximally unpredictable for the speaker, whereas the answer to the latter was expected to be maximally predictable. The experimental material included a range of conditions which were expected to elicit extreme as well as intermediate scores. Thus, on the one hand there were questions with a confirmative character, i.e. they contained the particle *dus* ('so') and/or the utterance-final tag *hè?* (', right?'). Such questions were hypothesised to produce high scores on a predictability scale[25]. On the other hand, answers to yes-no questions were expected to be maximally unpredictable: for the speaker, 'yes' or 'no' would seem equally probable (e.g. Kiefer 1981). Accordingly, such questions were expected to elicit low scores on a predictability scale. In addition, the questions were systematically varied with respect to the absence vs. presence of negation. This was also done with a view to creating intermediate categories, considering that negation has been suggested to cause an information question to become more like a confirmation question (e.g., Belnap & Steel 1976:111). Accordingly, the answer to a negated question was expected to be more predictable for the questioner than the answer to its non-negated counterpart. In all, we meant to vary the material along the following parameters:

- [+ INVERSION] vs. [– INVERSION]
  (i.e. yes-no question vs. declarative question)
- [+ NEGATION] vs. [– NEGATION]
- [+ *DUS*] vs. [– *DUS*]
- [+ *HÈ?*] vs. [– *HÈ?*]

Prior to the experiment proper we checked the acceptability of all possible cross-combinations with four native speakers of Dutch. According to these, one combination had to be eliminated: yes-no questions were unanimously felt to be incompatible with the confirmative tag *hè?*. Thus, an utterance such as *Vertrekt de trein om zes uur, hè?* ('Does the train leave at six, right?') was judged ungrammatical. This incompatibility was taken as preliminary evidence that the yes-no question is suitable only for eliciting information, not confirmation (see also Droste 1972:127). It meant, however, that the design could not be fully orthogonal. Consequently, two analyses of variance were carried out, one including both YQ and DQ in eight conditions, and one including only DQ, however with the added condition [+*HÈ?*] vs. [–*HÈ?*]. Figure 5.11 illustrates the design.

[25] Note that, strictly speaking, *dus* is not confirmative. Rather, it indicates that the speaker feels justified in making an inference. This, in turn, is felt to increase the predictability of the response. *Hè?* more explicitly appeals for confirmation. Thus, in *dus* and *hè?* the predictability of answers would seem to spring from slightly different speaker attitudes.

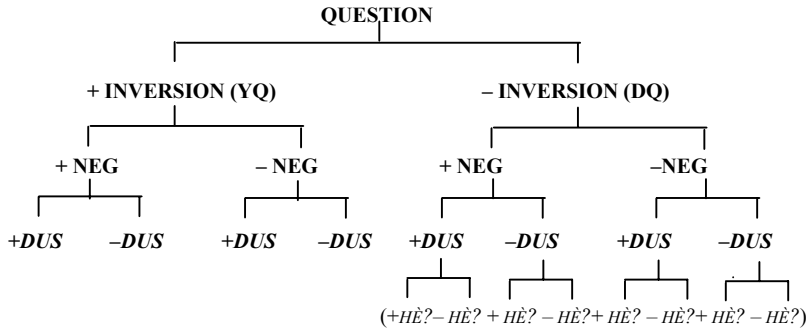


Figure 5.11. The experimental conditions.

Materials

Three declarative questions were made up to be used as test sentences: (a) *De laatste bus is al weg?*, (b) *Ze hebben mooi weer voorspeld?*, and (c) *Jullie kwamen op tijd aan?*[26]. Potentially, each of these could serve as a question for information the answer to which was maximally unpredictable for a questioner. These three core questions were then modified so as to feature INVERSION, NEGATION, and the particle *DUS*; the [–inversion] subset additionally featured *HÈ?*. This resulted in 3x12=36 test sentences. For one of the three core questions, part I of the Appendix illustrates the 12 conditions.

To avoid circularity, the experiment was carried out with written rather than with spoken questions. Since it sought to establish whether the observed intonational difference between YQ and DQ might have corresponded with a difference in communicative function, possible evidence of the latter had to be obtained independently. After randomisation, the 36 test sentences were presented in print to 32 native speakers of Dutch, all of them staff members or students of the universities of Nijmegen and Leiden. Primarily, subjects had to make a binary choice: was, for an imaginary speaker of the written question, the answer *unpredictable* or *predictable*? The former option obviously left no room for gradations. The latter, by contrast, was viewed as gradable, and subjects opting for it had to indicate the *degree* of predictablity. When the answer was judged maximally

[26]In English, these ran (a) *The last bus has already left?*, (b) *They have forecast fair weather?*, and (c) *You got there on time?*.

predictable, they ticked 100%. When the answer was felt to be less than 100% predictable, percentages of 80, 60, 40 and 20 were printed, but people were instructed to feel free to tick intermediate percentages. In addition, they had to indicate whether the answer was *predictably yes* or *predictably no* (recall that half of the questions featured negation). This division was introduced to make the task less monotonous for the subjects. Being not directly relevant to the research question, it is not taken into further account here[27]. Order-of-presentation effects were controlled for by varying the order of the questions (2x), as well as the orders of *maximally unpredictable* vs. *maximally predictable*, and of *predictably yes* vs. *predictably no* (4x); this resulted in eight different orders. A sample of the instruction form and excerpts of the score form are given in part II of the Appendix[28].

5.4.2 Results (1): YQ vs. DQ

The scores were subjected to an analysis of variance (SPSS Manova, repeated measures), with four within-subject factors: SENTENCE (3 levels), INVERSION (2 levels), NEGATION (2 levels) and *DUS* (2 levels). As in the analyses of variance in the chapters 3 and 4, the significance level was set at .01. All factors were regarded as fixed[29]. Table 5.1 presents the Huynh-Feldt adjusted p-values, together with the values for F and η^2. The significant interactions NEGATION x SENTENCE and INVERSION x NEGATION were ordinal, that is, differences between scores showed the same direction. The only disordinal significant interaction, *DUS* x NEGATION, is briefly discussed below. No significant higher-order interactions were found. Since there was no effect of sentence, scores were pooled across the three sentences.

[27] As a matter of fact, we found that, in 95.4% of the questions whose answers were judged predictable, positive questions correlated with 'predictably yes', negative questions with 'predictably no'. As could be expected, predictability of the answer included the polarity as presented by the questioner.

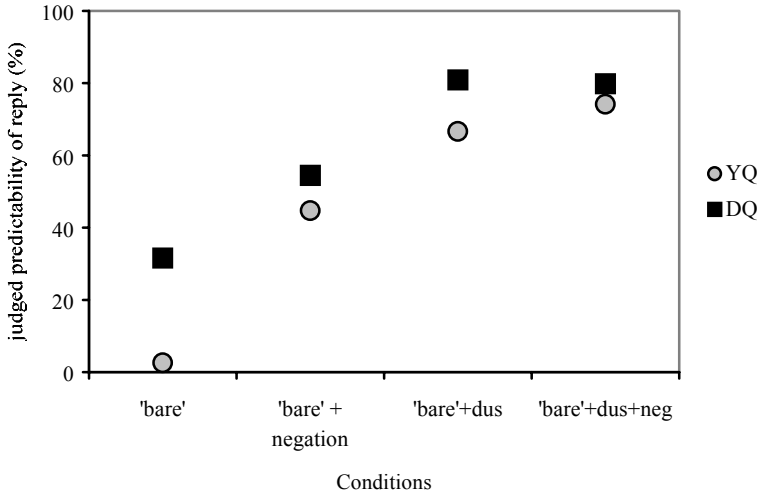
[28] After the subjects had ticked percentages, they were presented with a fourth question in the same twelve conditions as the three questions in the main task. This time, they had to order the different versions according to the predictability of the answers. This supplementary task was meant to act as a consistency check. As the results fully agreed with the scores on the main task, we will not take them into further consideration.

[29] Although the three experimental sentences were not randomly chosen they can yet be treated as random, given that they could have been replaced by other equally acceptable sentences without affecting the research question or conclusions (cf. Jackson & Brashers 1994).

Table 5.1. Main effects of four factors on predictability scores, together with the significant second-order interactions. Significance level: .01.

Factor/Interaction	df1, df2	F	p	eta square[30] (η^2)
SENTENCE	2, 62	4.44	.017	.038
INVERSION	1, 31	51.90	<.001	.626
NEGATION	1, 31	101.58	<.001	.766
DUS	1, 31	311.17	<.001	.909
SENTENCE × NEGATION	2, 62	7.99	.001	.175
INVERSION × NEGATION	1, 31	21.48	<.000	.409
NEGATION × DUS	1, 31	118.42	<.000	.792

Figure 5.12 plots the main effects as a function of question type.



[30] Note that eta square represents the default index for effect strength provided by SPSS, not the contribution of the factor to the total sums of squares (see also chapter 3, § 3.5).

Figure 5.12. Mean percentages of judged predictability of the reply as a function of the factors INVERSION, NEGATION, *DUS*. Each mean value represents 96 scores.

As the table shows, the subjects significantly differentiated between the questions whose only difference was in [+INVERSION] vs. [–INVERSION] (i.e. 'bare' YQ vs. 'bare' DQ, respectively). The very low mean predictability score on the yes-no questions (2.6%) provided clear evidence that replies to this question type were typically judged unpredictable, i.e. the answer might be equally *yes* or *no*. In its bare form, this question type scored '0' (i.e. 'maximally unpredictable') in 94% of the cases.

By contrast, the appreciably higher mean predictability score on declarative questions (31.6%) revealed that answers to this question type are, to some extent, predictable for the speaker; the difference with YQ was significant. At the same time, the scores on the predictability scale remained well below the maximum of '100% predictable'. Presumably, the additional features *DUS,* NEGATION and *HÈ?* acted as a ceiling, causing subjects to somewhat reduce their scores on the 'bare' DQ tokens. Interestingly, DQ scored '0' in 35 out of the 96 cases. That is, in 36%, replies to DQ were judged equally unpredictable as replies to YQ. Except for one case, subjects were consistent here: whenever they assigned '0' to DQ, they gave the same score to YQ. This implicational relationship suggests that these subjects may have viewed the two question types as functionally equivalent. However, what is of major importance here is that, overall, the difference between ('bare') DQ and ('bare') YQ was significant.

Addition of the particle *DUS* produced a very strong effect on the degree of predictability of the reply (see Table 5.1). As expected, in either question type its presence substantially increased the mean scores: in YQ, this rose to 67%, in DQ, to 81%. Proportionally, however, YQ was affected most (cf. Figure 5.12).

NEGATION, similarly, caused the answer to become more predictable, in both question types. However, its effect was substantially weaker than that of *DUS*. The significant interaction INVERSION × NEGATION was ordinal, that is, in either question type, NEGATION caused the scores to increase. By contrast, the significant interaction *DUS* × NEGATION was disordinal. That is, in DQ the combined effects of *DUS* and NEGATION produced slightly lower scores, when compared with the effect of *DUS* alone, whereas in YQ scores became higher. What Figure 5.12 also shows is that the overall mean scores for predictability did not exceed '80'. This may be taken to reflect that speakers remain conscious of the possibility that hearers will not in fact come up with the predicted reply.

5.4.3 Results (2): DQ

Unlike YQ, the question type DQ allowed addition of the tag *HÈ?*, which was expected to result in substantially higher scores on the predictability scale. For the set of declarative questions/*HÈ?* questions, we will use the term 'non-inverted questions' (NIQ). In order to judge the relative effect of *HÈ?* vis-à-vis the other factors, a separate analysis of variance was carried out (SPSS MANOVA, repeated measures), with four fixed within-subject factors: SENTENCE (3 levels), NEGATION (2 levels), *DUS* (2 levels) and *HÈ?* (2 levels). As before, the significance level was set at .01 and scores

were pooled across the three different sentences. Table 5.2 presents the (Huynh-Feldt adjusted) p-values, together with the values for F and η^2. Figure 5.13 plots the mean scores for the different conditions.

Table 5.2. Main effects of four factors on predictability scores in NIQ, together with the significant second and third order interactions. Significance level: .01.

Factor/Interaction	df1, df2	F	p	eta square (η^2)
SENTENCE	2, 62	1.10	.335	.032
NEGATION	1, 31	7.80	.009	.201
DUS	1, 31	99.10	<.000	.762
HÈ?	1, 31	112.65	<.000	.784
SENTENCE × NEGATION	2, 62	6.28	.004	.123
NEGATION × DUS	1, 31	24.40	<.000	.440
NEGATION × HÈ?	1, 31	50.06	<.000	.618
DUS × HÈ?	1, 31	194.00	<.000	.862
NEG × DUS × HÈ?	1, 31	17.48	<.000	.361

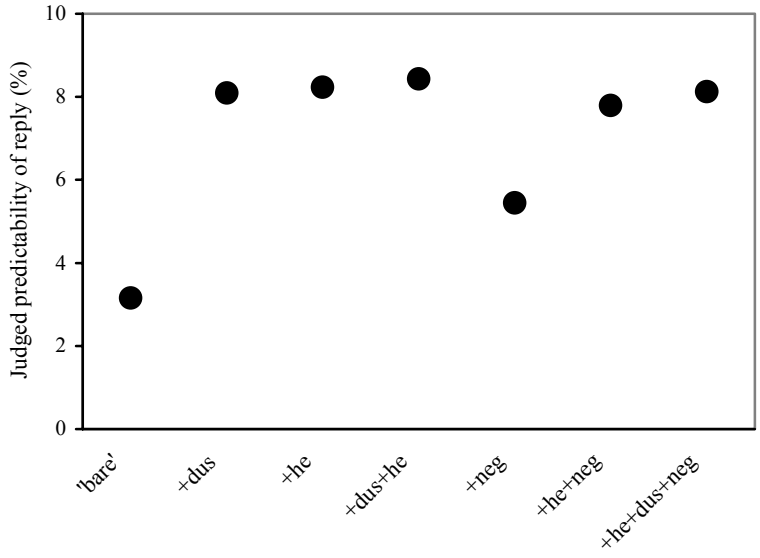


Figure 5.13. Mean percentages of judged predictability of the reply in NIQ (see text) as a function of the factors NEGATION, *DUS*, NEGATION +*DUS*. Each mean value represents 96 scores.

It is clear that *HÈ?* acted as a slightly stronger boost on the predictability scores than did *DUS.* When both occurred together, scores went up even further, producing the highest overall mean of 84.4. Combination of NEGATION with *HÈ?* and with *HÈ?+DUS* resulted in slightly lower scores, that is, *HÈ?* questions elicited the highest overall mean score of 84.4 with *DUS* but without NEGATION.

The three significant second-order interactions displayed roughly the same, slightly disordinal, patterns. The interaction DUS × HÈ? reflected that the presence of *hè?* acted as a very strong boost to the predictability scores when *dus* was lacking; however, when combined with *dus*, these scores hardly altered. Likewise, the interaction NEGATION × DUS indicated that the effect of negation was considerable as long as there was no *dus*, whereas a combination with *dus* brought about very little change in these scores. In a similar vein, the interaction NEGATION × HÈ? showed that, in the absence of *hè?,* negation caused the scores to rise, whereas in combination with *hè?*, scores became slightly lower. These tendencies were also apparent in the three-way significant interaction NEG × DUS × HÈ?. In all, the interactional patterns held a confirmation of the main effects: addition of *hè?*, *dus* or negation to 'bare' DQ caused a substantial increase in the mean scores, but combinations of either two of these variables did not truly alter the scores.

5.4.4 Discussion and conclusions

The aim of the present experiment was to establish whether yes-no questions and declarative questions differ in their communicative functions. While the former type is generally regarded as a true information question, it has been hinted in the literature that the latter might be better suited for asking confirmation. We assumed that the difference between 'information question' and 'confirmation question' can be expressed as the degree of predictability (for the speaker) of the corresponding reply. On a continuum from maximally unpredictable to maximally predictable, we expected yes-no questions to represent the base line in that, for a given speaker, the reply was maximally unpredictable. On the other hand, questions with the particle *DUS* and/or the tag *HÈ?* were expected to elicit very high scores for predictability. If the declarative question primarily functions as a confirmation question, it ought to get significantly higher predictability scores than the yes-no question or, better still, its scores ought to be close to those on questions with *DUS* and/or *HÈ?*. Likewise, we expected NEGATION to also elicit higher predictability scores, although our predictions on this variable were less concrete.

The results generally supported our expectations. First, subjects judged that, for a given speaker, the answer to a yes-no question is maximally unpredictable, i.e., *yes* is as likely as *no* (mean predictability score: 2.6%). Second, answers to DQ were judged significantly more predictable (mean predictability score 31.6%), indicating that the

communicative functions of YQ and DQ are not fully equivalent. In either question type, however, the particle *DUS* acted as a strong boost to the scores. Yet, scores on YQ remained systematically lower than on DQ, even in the presence of *DUS*. Apparently, inversion is inherently less compatible with a request for confirmation. This can also be inferred from the relatively low occurrence in YQ of the score '100' (indicating maximum predictability of the reply). Table 5.3 compares the incidence of '100' in YQ and DQ when *DUS* or *HÈ?* were present and could thus be expected to be strongly conducive to maximum scores.

Table 5.3. Incidence of the maximum predictability score '100' as a function of [± inversion] (i.e. YQ vs. DQ) with addition of ***DUS*** or ***HÈ?***. In each column, N=192 (questions with NEGATION were included).

Score	YQ + DUS	DQ + DUS	DQ + HÈ?
'100'	27 (14%)	58 (30%)	66 (34%)

Likewise, NEGATION systematically caused predictability scores to increase, be it to a lesser extent than did *DUS*. In the set of non-inverted questions (NIQ), the tag *HÈ?* had a similar effect on the scores as *DUS* in the set including both DQ and YQ.

Summing up, scores on the basic categories [+INVERSION] and [–INVERSION] indicated that YQ and DQ systematically differ in their communicative functions. Although the scores on DQ lay well below the scores elicited by *DUS*, *HÈ?* or NEGATION, it is clear that DQ may be used to seek confirmation as well as information. At the same time the scores reflected a functional distance between DQ and a hè?-question, with the latter functioning solely as a request for confirmation. Yet, in more than one-third of the cases DQ still scored '0', i.e. 'maximally unpredictable'; in fact, the intermediate level of the mean scores for DQ obviously derived from this relatively frequent occurrence of '0' scores. These scores were not randomly distributed, however. While roughly one half of the subjects gave '0' scores to both YQ and DQ, the other half clearly differentiated between the two, restricting the '0' score to YQ. Obviously, this finding is open to more than one interpretation. First, the former set of subjects may have lumped together the two question types, simply because they failed to notice the subtle difference (in print) between [+ INVERSION] and [– INVERSION]. Alternatively, the fairly systematic distribution.of '0' scores may reflect that Dutch subjects have different intuitions about YQ vs. DQ. That is, for some of them DQ can be used to gain information as well as confirmation, whereas for others YQ must be used to seek information, DQ for confirmation. This interpretation would be in line with informal observations that, in spontaneous speech, some speakers make frequent use of this question type whereas others never do. By the same token, it seems that some listeners persistently fail to interpret a declarative question as a *question*, forcing the speaker to repeat the utterance and to make its interrogative status explicit by other means.

The upshot of the experiment is that we are now able to answer the question that led to the experiment in the first place. It seems reasonable to assume that DQ's (minor) phonetic departures from the Frequency Code were motivated by the somewhat hybrid nature of this question type: it is fit for seeking information as well as

confirmation. Given that, in the latter function, DQ is a less prototypical question (cf. chapter 1, § 1.3.1), one may expect a (slightly) less prototypical intonation. It must be emphasised, however, that, on aggregate, the acoustic results on DQ confirmed our prediction (based, in part, on the Frequency Code) that DQ would display a larger amount of high(er) question pitch than WQ or YQ.

5.5 Summary

The present chapter further extended the analysis of the (question) intonation phenomena reported in the chapters 3 and 4. These phenomena could be captured by the formula Q=H, expressing that the questions typically involved higher pitch levels than the corresponding statements (§ 5.1). This finding, together with the finding that pitch was higher as questionhood depended more strongly on intonation (i.e., Q=H followed the order WQ<YQ<DQ), suggested a connection with the non-linguistic biological Frequency Code (§ 5.2). This code has been claimed to account for the universal form-function relation between high/rising pitch and questions, and between low/falling pitch and statements (Ohala 1983, 1984). Although the notion of universal meaning is not uncontroversial in intonation theory, Gussenhoven (2002 to appear) points out that the effects of universal innate codes may, at times, be too pervasive to be ignored. Thus, biological codes may counteract language-specific changes that have resulted in 'unnatural' phonological forms, such as questions with final falling pitch, or statements with final rising pitch. Given that for the realisation of phonological form a certain amount of phonetic 'space' is available, speakers are free to use this space to convey meanings deriving from universal, innate codes. One of these codes is the Frequency Code. Another is the Effort Code, which associates pitch excursion size with communicative explicitness, such that larger movements signal greater significance and emphasis (Gussenhoven 2002 to appear). Our findings on Dutch question intonation appeared to tie in quite well with Gussenhoven's views on the impact and operation of universal codes. Thus, the results made it plausible for high question pitch to have been motivated, to a considerable extent, by the non-linguistic Frequency Code and to have arisen in the phonetic implementation. However, we concluded that this code could also manifest itself in the phonological domain. The *phonological* reflexes of the FC in Q=H were:

- **H%** (final rise). This powerful prosodic attribute of interrogativity was viewed as a phonologisation of the Frequency Code. Assuming that pitch range variation was, by origin, gradient, discrete contrasts may have developed in the course of time and become part of the phonological system (cf. Gussenhoven 1999b). Although conventionalised H% is no longer a direct expression of the FC, the FC may yet assert itself indirectly by making speakers select this particular category.

- **H*** (high plateau). The various high plateaus in the questions increased the overall quantity of high pitch. Although it proved possible to capture all plateau types in a single phonological adjustment rule DELETION (an extension of the rule proposed by Grabe 1998), it was not possible to regard this rule as a

phonologisation of the Frequency Code, given that it also produced some low plateaus in questions, as well as some high plateaus in statements. Instead, we looked upon the question plateaus as products of *ad hoc* decisions to delete L tones with a view to raising overall pitch. The decision to apply such a phonological rule would then still be motivated by the FC. Evidence for this analysis came from results of a perception experiment (§ 5.3.2.1), as well as from the distribution of the H* H% plateau across the three question types (§ 5.3.2.2).

Phonetic reflexes of the Frequency Code in Q=H were found in:

- **The implementation of the initial boundary tone %L.** In the questions, mean f0 values of this category were higher than in the statements (§ 5.3.5).

- **The implementation of the final boundary tone H%.** Although the excursions of this category appeared fairly constant across the three question types, they were realised in successively higher portions of the range in the order WQ<YQ<DQ. This was due to corresponding increases in the values of the rises' onsets and offsets (PL's) (§ 5.3.1).

- **The f0 minima**. In the questions, mean minimum f0 values were higher than in the statements (§ 5.3.3).

In sum, over and above their categorical choices which, in part, may have been motivated by the FC, speakers also exploited the phonetic space to give (further) expression to universal meanings deriving from this FC.

The ensemble of raised utterance onset, raised minimum f0 and raised PL (postnuclear low) was taken to reflect a globally raised register level. This upward shift throws the question into prosodic relief, which probably serves to alert the listener to the fairly specific discoursal function of the utterance. Manipulation of register level (upward as well as downward) occurs, notably, in the area of attitudes and emotions and can be viewed as a general, gradable device for expressing or supporting various abstract meanings deriving from universal codes.

What could *not* be explained on the basis of the Frequency Code were the relatively many asymmetric accent patterns observed in the questions. These patterns largely resulted from local differences in accent size/peak height, i.e. accents in the questions were either larger or smaller than in the statements. It seemed likely that this cross-categorial variation reflected differences in focus. This, in turn, suggested the possible operation of the Effort Code, another biological code whose universal meanings may, likewise, be conveyed in the phonetic implementation. However, further exploration of this code's possible effects on the data requires a better understanding of the intricacies of focus in questions. For this reason, further analysis of the asymmetric accent patterns had to be deferred to chapter 6, which addresses this issue.

On two counts, f0 values in the declarative question type (DQ) slightly deviated from

predictions made on the basis of the FC. In order to establish wheter this might have been caused by a functional difference between the question types YQ and DQ, we carried out a pen-and-paper experiment (§ 5.4 and subsections). The results indicated that DQ is functionally unlike YQ in that it may be used to elicit confirmation as well as information. This dual function was taken to account for DQ's slightly hybrid phonetic shape. Nonetheless, it should be borne in mind that, otherwise, pitch values in DQ perfectly agreed with predictions based on the FC.

It will be obvious that our analysis has a fairly strong universalist flavour. Across the board, the data analysed in the present chapter suggested that the use of high/rising pitch in the questions was largely non-arbitrary. Instead, we took it to derive from the innate Frequency Code, which may assert itself in the phonetics as well as in the phonology. Moreover, it is possible that a better understanding of focus in questions (chapter 6) will lead to the insight that the hitherto unexplained accent asymmetries in fact rested on another biological code, the Effort Code. It cannot be doubted that, in diachronic processes, universal extralinguistic codes have left their marks on linguistic systems: phonological categories may in fact be discretised bits of universal meaning. That is, though intonation is essentially linguistically structured, this structure may (have) be(en) shaped to some extent by the operation of innate codes. Synchronically, these codes continue to exert their influence on speech, in the production as well as in the perception domain. Thus, we saw that they may affect speakers' choices from the inventory of phonological categories. More typically, however, the universal codes operate in the phonetic domain, where their effects are more immediate. In all, our analysis reconciles the well-established view that linguistic signs are, typically, arbitrary with the universalist perspective. Hence, the following comment by Ultan may appropriately conclude the present chapter:

> It is easy to see why any kind of sound symbolism must always be of a peripheral nature. If it were not, the extremely limited number of contrastive phonological features available in any language would hardly suffice to represent the enormous complexity of all the semantic distinctions necessary in human communication. Nevertheless, the basic and inescapable principle of the arbitrariness of language symbols is neither absolute nor inviolable. (Ultan, 1970:551)

To this it may be added that, if there is a part of the grammar that would seem more informed by iconicity than any other, it is no doubt question intonation.

Chapter 6

Questions, accent and focus

6.1 Introduction.

One of the main findings in chapter 4 was that, in the question material, there was a strong tendency towards asymmetric accent patterns, whereas in the statements accents were broadly equal. The yes-no questions (YQ) and declarative questions (DQ) featured a reduced subject accent and raised object accent. In the wh-questions (WQ), the mean peak f0 of the initial accent on the wh-word was higher than the mean peak f0 of the subject and object accents, if the latter were at all present (only 12% of the subjects and 63% of the objects carried accent). Hence, there were differences in relative accent scaling *between* the discoursal categories 'statement' and 'question' (ST vs. Q), as well as *within* the category 'question' (WQ vs. YQ/DQ). Towards the end of chapter 5, the question arose whether these asymmetric accent patterns might have been reflections of underlying, non-neutral focus structures. The present chapter aims to answer this question. The research questions can be formulated as follows:

- Does the discoursal difference between statement and question correspond with a difference in the *function* of focus?
- If there is such a difference, does it account for the different accent patterns observed in the statements and the questions?

The chapter is organised as follows. The first part is generally concerned with the relationship between accent patterns and focus (§ 6.2 and subsections). Looking, among other things, at the various motivations that may underlie focus, we distinguish three focus types. Section 6.2.4 proposes to identify the domains, over which focus has scope, with the pragmatic constituents topic and comment. Consequences of this proposal for accent patterns in yes-no questions and wh-questions are discussed in §§ 6.2.5-7. Section 6.2.8 indicates that an approach in terms of topic-comment may, also, be beneficial to the analysis of accent patterns in *statements*. This part of the chapter then concludes with an overview of acoustic properties of contrastive focus, on the assumption that these features may have given rise to the asymmetric accent patterns in the questions (§ 6.2.9). Next, on the basis of the preceding sections, predictions are made of the accent patterns in the four experimental utterance types. These predictions are then compared with the patterns as they were actually observed (§ 6.3 and subsections). In § 6.3.5, the results are discussed and related to the biological Effort Code. The chapter concludes with a summary in § 6.4.

6.2 Accent and focus

To a large extent, the shape of an intonation contour is determined by the position and realisation of the accents (e.g. Bolinger 1972a; Liberman & Pierrehumbert 1984; Eady & Cooper 1986). These, in turn, rest on a certain focus structure, that is, a speaker decides to call the listener's attention to a particular entity by bringing this in focus and giving it an accent (e.g. Eady *et al.* 1986; Terken & Nooteboom 1987; Ladd & Terken 1995; Rump & Collier 1996). On his part, a listener makes use of such prosodic information to decode the intended focus structure: "[…] when listeners have determined where accent falls, they have located the focused or informationally prominent part of an utterance: an active search for accent may therefore represent an active search for the semantically most central portion of a speaker's message" (Cutler, Dahan & Van Donselaar 1997:174).

Under the so-called Focus-to-Accent view (FTA, Gussenhoven 1985:125, for discussion see Ladd 1996:163), the relationship between focus and accented item is indirect rather than direct. That is, while a speaker is obviously free to choose the entity he wishes to focus on, he relies on independent principles to determine which word/syllable within that entity is going to carry the pitch accent. In several quarters, these principles are claimed to be chiefly semantically motivated. For instance, Gussenhoven's (1984a) Sentence Accent Assignment Rule (SAAR, see further § 6.2.2 below) claims that (i) the semantic constituents Argument and Predicate, when adjacent, may merge to form a single focus domain, and (ii) that within this composite domain accent is carried by the Argument[1]. This implies that focus may extend over a domain that is larger than the accented word ('focus projection'[2]). Conversely, within a focused entity certain words may remain unaccented, notably the verb.

Many utterances have 'broad focus', that is, focus has scope over the entire utterance or at least over some multi-word constituent within it (cf. Ladd 1980, 1996:161). This is not to say that speakers may not have specific reasons for narrowing this scope, for instance by exclusively focusing on the verb (*I want to BUY a house (rather than rent it)*), on a lexically non-stressed syllable (*I meant SchuBERT, not SchuMANN*), or on a non-content word (*I said THE bank, not A bank*). However, in the more common case of broad focus, accent placement is widely assumed to obey certain structural principles.

Although the view that accentuation is mediated by focus structure would seem largely uncontroversial now[3], Bolinger (e.g. 1972b) has always been a strong advocate of a *direct* link between the word a speaker wishes to highlight, and the accent. This view is explicitly refuted by Ladd (1987, 1996:161ff), among other things on the evidence of crosslinguistic accentuation patterns in *questions*. Though acknowledging that "Accentuation has not been studied in questions anywhere near as much as in statements […]" (p. 174), Ladd points out that, in some languages, the

[1] That semantic argument structure can determine the location of sentence accents has also been proposed by Schmerling (1976) and Selkirk (1984).
[2] The term is Chomsky's, cf. Gussenhoven (1999a).
[3] But see Hoekstra (2000), whose algorithm for the assignment of Dutch sentence accents (developed as part of a text-to-speech system) dispenses with the notion of focus.

pattern of accentuation seems to be affected by the opposition statement/question. Thus, in neutral Russian statements it is the noun that is accented whereas in corresponding neutral yes-no questions, it is the verb (cf. Ladd 1996:168-167). This, according to Ladd, suggests that there is no universal algorithm that directly maps accents onto discoursally important *words*. Ladd goes on to say that Bolinger's view also predicts that, cross-linguistically, the wh-word in wh-questions is accented, given that it is the most important word of the utterance. But while this is true in languages like Romanian and Hungarian, in Italian and Portuguese it seems doubtful whether the wh-word carries the most prominent accent. Observations such as these lead Ladd to conclude that accentuation cannot be prompted by a universal principle according to which a speaker directly highlights the words that are important to the discourse. Rather, he speculates, there seem to be language-specific differences, stemming from differences between grammars.

Whereas Ladd's crosslinguistic examples thus indicate that, between (one-accent) statements and questions, accent *placement* may sometimes vary, our Dutch (two-accent) statements and questions systematically differed in the *scaling* of their accents. Thus, while accent patterns in the statements were usually symmetric in that both accents were of roughly equal size, the questions mostly displayed asymmetries between their accents. In chapter 5, we tentatively attributed this variability in accent patterning to differences in focus structure (§ 5.3.4.1). However, it was evident that further conclusions as regards this issue require a more thorough understanding of focus in questions. As a preliminary, the next section considers different motives for bringing constituents in focus, that is, it identifies different *types* of focus.

6.2.1 Motives underlying focus ('focus types')

As noted above a speaker bringing an entity in focus does so with a view to drawing an addressee's particular attention to it; intonationally, he achieves this by realising a pitch accent. It is clear that a speaker may have different reasons for focusing, the most common being that the focused entity introduces information that the speaker assumes is new to the discourse or to the interlocutors' common background. Typically, the information highlighted by this type of focus comes from a set of possibilities that is potentially unlimited (cf. Chafe 1976:34). Thus, in utterance (6.1a), the [+F] information is assumed to be new to the listener; square brackets reflect pragmatic constituency (see further below), +F indicates focus, –F non-focus, and accented words are capitalised.

(6.1a) $[\text{We}]_{-F}$ $[\text{'re going to need a CHAIRMAN}]_{+F}$

Considering that, in (6.1a), the [+focus] part adds new information to the interlocutors' background, we might use the term FOCUS OF ADDITION (cf. Gussenhoven 1984a). However, discussions in the literature have made it clear that 'newness' is not the sole motive for bringing items in focus (e.g. Kiss 1998). In fact, it seems useful to distinguish more types of focus, among other things because the phonetic correlates of these may, also, differ. Accordingly, Chafe (1976:35)

distinguishes FOCUS OF CONTRAST, which communicates "[…] that a certain focus item rather than other possible ones is correct […]". According to Chafe, FOCUS OF CONTRAST is present whenever the focused part can be interpreted as being implicitly followed by phrases such as *rather than*, or *instead of*. This focus type is illustrated in (6.1b) and (6.1c) below.

(6.1b) [Paul's going to be]$_{-F}$ [CHAIRMAN]$_{+F}$ (rather than secretary)
(6.1c) [PAUL]$_{+F}$ (rather than Peter) ['s going to be chairman]$_{-F}$

Throughout the present chapter, the term FOCUS OF CONTRAST will be tied to this essential notion of 'correct item from a subset'[4]. A non-prosodic device for marking items for contrastive focus is clefting (cf. Kiefer 1980:106; Wunderlich 1981:140; Quirk *et al.* 1987[5]; Kiss 1998). Thus, (6.1b,c) could be clefted to form *It's CHAIRMAN Paul is going to be*, and *It's PAUL who's going to be chairman*, respectively.

Crucially, the two focus types are fully independent in that FOCUS OF CONTRAST may involve new as well as given items. Thus in (6.1b), uttered in the course of a discussion about the distribution of administrative functions, *chairman* is probably given rather than new. However, it is implicitly contrasted here with a set of other possible functions out of which only *chairman* is correct. The same holds for (6.1c), where it is perfectly possible for *Paul* to be at the same time given as well as the correct person from an implicit set of candidates. Finally, consider (6.1d), which has two accents rather than one.

(6.1d) [PAUL]$_{+F}$ ['s going to be CHAIRMAN]$_{+F}$

By explicitly bringing *Paul* in focus here, the speaker causes the hearer to (re)activate the referent of *Paul* with a view to (re)introducing it to the current discourse. Obviously, this motivation for focusing differs from the two exemplified above; it may be termed FOCUS OF INTRODUCTION. Note that a division of focus into subtypes should not be taken to suggest that focus is no more than a convenient cover term for what is in fact a heterogeneous lot. Across the three types distinguished here, there is a strong common denominator in that focus always seeks to call the hearer's attention to a particular item; what is variable are the speaker's reasons for doing so[6]. As is apparent from the above examples, the scope of focus may extend over a group of words ('focus projection'); within such a focus domain,

[4] When we mention other terms used in the literature, such as 'contrastive focus', 'contrastiveness' etc., this does not necessarily mean that these also include this notion of 'correctness'.
[5] "The highlighted element has the full implication of contrastive focus: the rest of the clause is taken as given, and a contrast is inferred with other items which might have filled the focal or 'hinge' position in the sentence" (Quirk *et al.* 1987:951).
[6] See also Dik *et al*'s (1980) more extensive typology of focus. Gundel (1999), likewise, proposes different kinds of focus. Her 'contrastive focus' would seem to include both the above FOCUS OF CONTRAST and FOCUS OF INTRODUCTION, whereas her 'psychological focus' is explicitly discoursal in that it refers to "entities in a discourse that both speaker's and addressee's attention is currently focused on" and that are therefore likely to be continued as topics of subsequent utterances. In this latter sense, however, focus is conceptually different from the sense in which it is taken in the present study (for the terminological confusion as regards the term 'focus', see e.g. Ladd 1996:293 footnote 3; Gundel 1999:302)

accent has to be assigned. The next section provides a brief outline of SAAR (Sentence Accent Assignment Rule), the algorithm for the formation of focus domains and the assignment of accents as proposed by Gussenhoven (1984a, 1992, 1999a).

6.2.2 Focus and accentuation according to SAAR

In Gussenhoven (1984a:18), focus is defined as a binary feature with [+focus] marking "the speaker's declared contribution to the conversation, whereas [–focus] constitutes his cognitive starting point". This contribution modifies the hearer's discourse background, that is, the content of the focused part causes the hearer to update his current state of information. It is assumed that utterances are always marked for focus and that focus involves semantic rather than syntactic constituents.

While it is essentially unpredictable which part a speaker is going to mark as [+focus] or [–focus], the location of the pitch accent *inside* this part can be predicted by the Sentence Accent Assignment Rule (SAAR), which is claimed to apply in both English and Dutch (Gussenhoven 1984a:32). The rule distinguishes the three major semantic constituents Arguments (A; figuring as subjects and objects), Predicates (P; figuring as syntactic predicates), and Conditions[7] (C; figuring as adverbials). Whenever a focussed Argument and Predicate are adjacent (disregarding any intervening non-focused Argument or Condition), they merge into a single focus domain. Within this domain, the accent is automatically assigned to the Argument. Thus, in [*Our dog*]$_{A}$ [*has disappeared*]$_{P,}$ A and P merge into the [+focus] domain [*Our DOG has disappeared*], and the Argument *dog* carries the accent. Although, within such an extended focus domain, the Predicate is barred from being accented, it is yet [+focus]. In the word selected for accentuation, accent will commonly fall on the syllable that bears lexical stress[8].

6.2.3 Focus in questions

So far, most work on focus has concerned statements; the issue of focus in *questions* has received relatively scant attention. In the present study, questions are explicitly viewed as utterances in their own right (cf. chapter 1): the assumption is that a question's focus structure and the corresponding accent pattern do not depend on the actual content of a subsequent answer. That is, we approach questions from a pragmatic perspective and assume that a speaker tries to communicate as clearly as possible which part of the utterance is the object of questioning, and which is not. To accomplish this, he decides on a focus structure, which is then conveyed to the listener by means of a certain accent pattern. In the sections below, we try to answer the question whether the function of focus in questions differs from that in

[7] Later renamed Modifiers (Gussenhoven 1992, 1999a).
[8] According to a later formulation of SAAR, *all* accentable syllables start off being accented. Following this, a deletion rule causes Arguments, Predicates and Modifiers outside focused constituents to be deaccented. The same happens to any focused P adjacent to an accented A (intervening non-focused A or M may be disregarded); see Gussenhoven 1999a:45.

statements. Prior to this, however, we introduce a pragmatic distinction which is going to play a crucial role in our analysis.

6.2.4 Topic-comment structure

At the level of pragmatics, a statement may be divided into two parts, the part functioning as the 'topic' (which refers to the person or thing about which something is said), and the part functioning as the 'comment' (which adds something about that topic)[9]. The topic-comment dichotomy is assumed to be a linguistic universal (cf. Gundel 1988). As Gundel (1977:16) formulates it, "[…] there is a good deal of formal evidence in various languages that native speakers distinguish 'what the sentence is about' from the rest of the sentence and that adequate descriptions of languages must therefore take this notion into account." Although the topic will often coincide with the syntactic subject and occupy an initial position, this is not necessarily so. The comment may consist of the predicate but also of a noun phrase, a prepositional phrase etc. Considering that statements always seek to make some contribution to a given background, dividing them into the functional complements topic and comment seems an obvious and effective move.

At first blush, it may seem plausible for topic and comment to correspond with a [–focus] and a [+focus] domain, respectively. Considering that it is usually the comment that prompts the speaker to produce the statement in the first place, this comment will always be [+focus] (Gundel 1988); this is then FOCUS OF ADDITION. The topic, however, is not inevitably [–focus]. Although, quite often, it is already part of the interlocutors' background, the speaker may yet have reason to assume that it has to be explicitly (re)introduced to the current discourse. In that case, the topic is [+focus], too[10] (FOCUS OF INTRODUCTION). To give an example, a topic realised by a personal pronoun usually refers to someone who is already part of the discourse and can therefore be left out of focus. A topic realised by a proper name, however, is more likely to be new to the discourse and hence [+focus]. In sum, in a topic-comment articulation an utterance such as *Paul's going to be chairman* consists of two focus domains: the topic *Paul*, which may be either [–focus] or [+focus] depending on its status in the discourse, and the comment *'s going to be chairman*, which is obligatorily [+focus].

Below, we will argue that the focus type in comments of *questions* is FOCUS OF CONTRAST rather than FOCUS OF ADDITION. To make this clear, it is important to stress that comments in questions would appear to differ from comments in statements, in that they reflect presuppositions entertained by the speaker (which is why the speaker asks the question in the first place). As Higginbotham (1995:375) notes, "Acts of asking questions carry presuppositions, sometimes present in the situation of the utterance without being linguistically triggered, and sometimes present in the form of the utterance itself"[11]. What happens in a question is that the

[9] According to Gundel (1977:19), it was Hockett who, in 1958, first introduced the terms 'topic' and 'comment'.

[10] That a topic may be accented, too, has been suggested by, e.g., Schmerling (1976:94), Dik (1980) and Nooteboom & Kruyt (1987).

[11] Presuppositions can be semantic as well as pragmatic. Broadly, one could say that semantics primarily approaches presuppositions in terms of truth relations, whereas in a pragmatic approach presuppositions are assumed to depend on overall discourse context.

speaker brings a presupposed comment in focus in order to check whether it is the *correct* comment vis-à-vis other potential comments. This implication of 'correct item from a subset' makes FOCUS OF CONTRAST the appropriate focus type. In the literature, contrastively focused items have been associated with specific phonetic correlates. For instance, Brown *et al.* (1980:29) note: "In our data, speakers appear to distinguish not only between given and new, but also between new and contrast, in that what is contrasted regularly appears on a higher pitch than what is introduced as new which, in turn, appears on a higher pitch than what is introduced as given." Possibly these phonetic differences, associated with different focus types, account for the observed variation in accent scaling in our data. Therefore, the next sections explore how topic-comment articulation and distribution of focus types work out in yes-no questions and wh-questions. (Note that the presence/absence of inversion is irrelevant for the topic-comment division; what is said about yes-no questions also holds for declarative questions.)

6.2.5 Yes-no questions

We propose that, in yes-no questions, the comment represents the presupposition currently entertained by the speaker. Being the object of questioning, this comment also reflects that there is a set of related comments that might, potentially, also apply to the topic in hand. This set is inherently constrained, though. A questioner asking whether Paul is going to be chairman does not intend to check non-related presuppositions, such as whether Paul may be going on a cycling holiday, or whether Paul is getting married next week. Indeed, replies to that effect would generally be infelicitous. From this naturally limited range, the questioner proffers his presupposition as the one he judges, provisionally, to be correct. By doing so, he implicitly contrasts it with other potentially correct presuppositions (one of which is the negation of the comment that is being put forward). Since the motivation for bringing the comment in focus is the wish to establish the latter's correctness vis-à-vis that of other members of a set of comments, we have FOCUS OF CONTRAST here[12] (cf. Chafe 1976:35, and § 6.2.1 above). For an example, let us consider the yes-no question version of sentence (6.1) above:

(6.1e) Is Paul going to be chairman?

As noted earlier, a well-tried device for determining the topic consists of testing whether phrases such as *As for..*, or *About…* can be felicitously inserted before the presumed topic (cf. Chafe 1967, Gundel 1977). This yields *Paul* for the topic function and, assuming that topic and comment are complementary, *is going to be chairman?* for the comment. Since the comment is the part of the sentence the speaker primarily wants to call the listener's attention to, it is [+focus]; we assume, for the present, that the topic is [–focus]. Clearly, the comment here is unlike its counterpart

[12] For isolated hints of some relation between questions and contrastiveness, see Van Es (1932; cf. chapter 2 § 2.2.3), Kiefer (1980), and Eady & Cooper (1986). The latter authors, having investigated accent pitch in questions, wonder "whether intonational emphasis on a particular constituent serves as an aid to comprehension in answering a question pertaining to that constituent" (p. 413).

in the statement version, given that its function is not to *add* something to the topic. Instead, the comment reflects the presupposition that prompted the speaker to utter the question, i.e., his 'cognitive starting point'. This presupposition is at the same time questioned, that is, its correctness is checked. The entire question may be paraphrased as 'At present, I believe that something can be said about Paul, viz. that he *is going to be chairman*, and I want to check whether that is correct'[13]. As the motive underlying focus is to establish the correctness of the presupposed comment alongside other potentially relevant comments pertaining to *Paul*, we have FOCUS OF CONTRAST here. According to SAAR, the comment's Argument (i.e. *chairman*) receives an accent.

As observed earlier, there may be reasons for the speaker to suppose that the referent of *Paul* is **not** part of the listener's current awareness. Consequently, the topic needs to be activated by bringing it in focus (FOCUS OF INTRODUCTION) and giving it an accent. In such cases, both constituents are [+focus] and carry an accent, be it for different reasons. To sum up, yes-no questions have an obligatory accent on the comment (FOCUS OF CONTRAST) and an optional accent on the topic (FOCUS OF INTRODUCTION).

6.2.6 Wh-questions

In wh-questions, matters are slightly but not essentially different. In the question *Who's going to appoint the new chairman?*, the speaker entertains the (semantic) presupposition *Someone is going to appoint the new chairman* (e.g. Kiefer 1980:101), irrespective of whether this is actually the case. Given that this proposition is already part of the common background, it functions as the question's topic ('As for the person going to appoint the new chairman,…'). What is being questioned is the referent of the interrogative pronoun *Who*: this wh-phrase then functions as the comment (cf. Bolinger 1990). It presents the listener with a syntactic category defining the set of possible replies, in this case the set of persons that can be supposed relevant in the given context. The question may be paraphrased as: 'I believe that our presupposed information (i.e. that someone is going to appoint the new chairman) can be extended to include the person who will actually be the one to appoint this new chairman'. The comment's lexis further indicates that the speaker wants to be provided with the correct item from some set of relevant persons; accordingly, the focus type is FOCUS OF CONTRAST[14]. That the comment in wh-questions occupies the leftmost position illustrates that this position does not

[13] Superficially, the above paraphrase might seem to suggest that the questioner merely asks for *confirmation.* There is no inconsistency here, though. In chapter 5, results of a pen-and-paper experiment made it clear that yes-no questions are not confirmation questions, in that answers are essentially unpredictable for the speaker: they may be *yes* as well as *no.* When a questioner questions some comment he entertains, this is not to say that he can *predict* the reply.

[14] Note that this would not seem to apply to so-called 'open wh-questions'. An open question such as *What are questions?* does not, implicitly or explicitly, suggest a restricted set of possible comments. Rather, according to Groenendijk & Stokhof (1996:1108), there is an unlimited amount of answers that can be given, all of which may be appropriate and not necessarily mutually inconsistent. Since the open question does not determine a number of pre-set possible answers, replies have to be created 'from scratch'. However, as there is no consensus as to whether or not open questions are conceptually different from what may be called 'informative questions', we do not pursue this matter here.

automatically belong to the topic (cf. Gundel 1977:30)[15]. In sum, the primary function of the question word is to make the hearer select the correct item from a set of potential candidates.

As in yes-no questions, the topic of a wh-question may be [+focus] as well as [–focus]. When uttering *Who's going to appoint the new chairman?*, the speaker may be uncertain as to whether the implied topic sentence *Someone is going to appoint the new chairman* **is** actually part of the listener's background. Hence, he may want to explicitly (re)introduce it to the discourse. Accordingly, he projects two [+focus] domains, one for the topic, and one for the comment. In the topic sentence, the focus type is FOCUS OF INTRODUCTION; in the comment, it is FOCUS OF CONTRAST. Thus, both in yes-no questions and in wh-questions speakers are allowed two options: a compulsory accent on the comment with or without an accent on the topic.

6.2.7 Discussion

In the account proposed here, the generalisation is that yes-no questions, declarative questions and wh-questions are requests for the correct item from some explicit or implicit set. This analysis presupposes that the utterance is partitioned into the universal pragmatic constituents topic and comment. Comments in questions are claimed to have a different function from comments in statements, in that they involve a speaker's presuppositions together with the crucial notion 'correct member of a set'. In yes-no questions and declarative questions, the speaker checks whether the comment put forward by him is the correct one out of an implicit set of potentially appropriate comments. One obvious member of this set is the negation of the proffered comment. In fact, the set may well hold just two members: affirmation and negation. In wh-questions, the comment embodied by the wh-phrase (e.g. *Who? Where? When? How?*) reflects a wider range of options. From this set, the listener is requested to supply the correct referent. In both question types the speaker's motivation for bringing the comment in focus crucially involves the notion 'correct member of a set'; the focus type, according to Chafe (1976), is FOCUS OF CONTRAST. As will be seen below, this focus type is associated with certain phonetic correlates, which may account for the asymmetric accent patterns observed in our question materials.

Obviously, the comment in a question may vary as to its scope. The [+focus] domain may include an Argument and a Predicate (e.g. *is going to appoint the new CHAIRMAN?)*, but it may also involve a single word or even a single syllable (e.g. *chairMAN?)*, depending on the set of comments implied by the speaker. That is, the

[15] By contrast, in Halliday (1967b:212) the topic ('theme' in his terminology, defined as 'what is being talked about') typically goes together with sentence-initial position. Accordingly, the wh-word in wh-questions corresponds with the topic rather than with the comment ('rheme'). By the same token, Halliday contends that the topic in yes-no questions is the sentence-initial auxiliary verb used in the inverted construction, e.g. *did*. In questions, he argues, the theme characteristically focuses on the unknown, i.e. there is something the speaker does not know and wants to know. However, this line of reasoning is rejected by Gundel (1977:31) who wonders: "[…], then why isn't the theme (i.e. the topic) of an assertion that there is something the speaker does know and wants to tell the hearer?". In the present study, the topic need not necessarily occur in initial position and the wh-word is identified with the comment rather than with the topic. Neither is a question-initial auxiliary verb such as *do* or *did* viewed as the automatic candidate for the topic.

comment and hence the domain of FOCUS OF CONTRAST may be progressively narrowed. Obviously, listeners will have to decide on the intended scope of the [+focus] comment. We assume that their interpretation is guided by the discourse context as well as by the relative size of the accent, as is also the case in statements (where the nuclear accent is typically ambiguous between broad and narrow focus[16]).

Support for the claim that the comment in a question involves the Chafian notion 'correct item from a subset' may be found in an observation by Gussenhoven (1984a:47) about the focus distribution in replies to yes-no questions. First, he considers the wh-question/answer pair given in (6.2).

(6.2) A: Who was born in Paris in 194SIX?
B: JOHN (was born in Paris in 1946)

That *JOHN* is [+focus] here is only to be expected, according to Gussenhoven, given that "answers to wh-questions have the focus on the requested bit of information" (1984a:47; see also Dik *et al.* 1980). What surprises him, however, is that in (6.3) below, the yes-no question and the corresponding answer have the **same** focus distribution; in lexically fully specified affirmative answers, this seems the neutral situation.

(6.3) A: Is he the GARDener?
B: (YES,) he's the GARDener

Assuming that the requested bit of information involves the sentence's polarity, Gussenhoven would expect 'polarity focus' here, as would be the case in the negative answer to this question: *(NO,) he ISn't the gardener*[17]. In the affirmative reply, this would then yield *(YES,) he IS the gardener*, with polarity focus on the operator.

Yet, when a question is seen as a request for the correct comment from some set of comments, B's reply could be regarded as providing just that: by repeating the comment it indicates that the latter is correct. Indeed, a cooperative speaker may echo the correct comment together with the topic to which it pertains, as in (6.2B), so as to leave no doubt that the two belong together. The same then applies to the negative reply *(NO), he ISn't the gardener*, reflecting that speaker A's presupposed comment was **not** correct. This reply could be paraphrased as 'As for your presupposition that he is the gardener (topic), this is not correct (comment)'. What constituted the comment for A (*is the GARDENER?*) has become the topic for B; hence, it is [–focus]. B's comment (his contribution to the common background) is to negate this topic, that is, he brings only its polarity in focus. What is important here is that, both in the positive and the negative reply, the [+focus] part corresponds with the correct comment. In reply (6.4B) to the

[16] See also the results of perception experiments by Brown *et al.* (1980; § 6.2.4) and by Bartels & Kingston (1994), reported in § 6.2.9 below. In the latter, listeners resolving the focus ambiguity between *Amanda [had a banana]*$_{+F}$ and *Amanda had a [banana]*$_{+F}$, apparently used peak height as their main cue.

[17] In Gussenhoven's model, the mode [counterassertive] allows sentences to have only their polarity in focus, as in *The house IS/ISn't for sale*, or, in Dutch, *Het huis is WEL/NIET te koop* (cf. 1984a:46). In such instances of 'minimal focus', focus involves less than one of the major semantic constituents (A, P or C) formally required by SAAR (1984a:45, 49).

full-focus question (6.4A) Gussenhoven (1984a:48) would, likewise, expect the operator to have polarity focus (i.e., *the volcanoes ARE dormant*). Instead, he notes, the reply features two separate [+focus] domains and the operator *are* remains unaffected:

(6.4) A: Are the volcanoes dormant?
B: Yes, [the VOLCANOES] [are DORMANT]

In point of fact, one might say that the replies in (6.3) and (6.4) closely resemble one other, in that both echo the correct comment **and** the corresponding topic of the preceding question, leaving no room for misunderstanding. The only difference is in the optionality of the plus vs. minus focus status of the topic. Thus, in (6.3) the topic *he* (an anaphoric word) is [–focus] and hence unaccented, whereas in (6.4) the topic *the volcanoes* is [+focus] and carries accent.

In sum, the generalisation seems to be that, in positive as well as in negative replies to *yes-no questions*, the correctness of the comment is [+focus]. When correct, the comment is repeated, when incorrect, its polarity is changed. This seems to support our argument that questioning implies a request for the correctness of the comment. Cooperative replies may further indicate whether the *ensemble* of topic and comment is correct, by repeating both. As regards replies to *wh-questions*, several authors have expressed the view that the *referent* of the wh-phrase is contrastively focused (e.g. Wunderlich 1981), suggesting that the focus type of the wh-word itself is FOCUS OF CONTRAST[18]. In fact, it may be interesting to investigate whether there is a similar tendency in replies to yes-no questions (such as (6.3B) and (6.4B) to mimic the phonetics of the (contrastive) accents in the preceding questions (see also Eady & Cooper's comment in footnote 12 above).

To the analysis proposed here, one might object that there are utterances that would seem topic-less, such as the statement *Someone came in*, or the question *Was there anybody?*. Such sentences need not pose any problems, though. Following Gundel (1977:34), we propose that in such cases the topic can be equated with the current situation, e.g. with the time or the place about which something is being asserted or questioned. That is, in these so-called 'all-comment presentational sentences' (cf. Gundel 1988), the topic element is not overtly expressed; since it is assumed to be recoverable from the immediate context, it is deleted from the surface.

Finally one might wonder whether, under the present analysis, the notion focus may not be dispensed with altogether, on the idea that it is simply substituted by the topic/comment dichotomy. In fact, Gundel (1977:45) remarks "[…] given an adequate theory of topic-comment structure no separate notion of focus needs to be accounted for in the grammar." Now that would be true if the comment automatically corresponded with [+focus], the topic with [–focus]. This, however, is not the case: in our proposal the topic may both be [–focus] and [+focus].

[18] An example of a relation between wh-phrase and contrastiveness which is structural rather than prosodic can be found in Hungarian. In this language, a contrastively focused item always takes a preverbal position, whereas items in what we have termed FOCUS OF ADDITION occur postverbally. Now, a Hungarian wh-phrase always occurs in preverbal position, and this commonly also holds for its referent in the corresponding answer (Kiss 1998).

Summarising the argument so far, we have suggested that the issue of focus in questions becomes more tractable after a preliminary analysis at the pragmatic level. For communicative reasons, sentences are intrinsically divided into the pragmatic constituents topic and comment, each of which constitutes a focus domain of its own. The latter is obligatorily [+focus]; the former may be [–focus] or [+focus], depending on whether or not the speaker feels the need to explicitly (re)introduce it to the discourse. Since focus can be differently motivated, we distinguish three focus types, one of which only applies to topics (FOCUS OF INTRODUCTION). Of the other two focus types, both applying to comments, FOCUS OF ADDITION is only relevant in statements, whereas FOCUS OF CONTRAST may be relevant both in statements and in questions. Indeed, for the comments of questions it is the only appropriate focus type. We emphasise that, in the approach outlined here, FOCUS OF CONTRAST is viewed as a specific focus *type* associated with a specific motivation for bringing the constituent in focus. Thus contrastiveness, interpreted here in the Chafian sense, is neither regarded as a bivalent semantic feature entirely independent of focus structure, nor as a gradient, derived property of focus itself (cf. Bartels & Kingston 1994). The diagram in Figure 6.1 sums up the analysis.

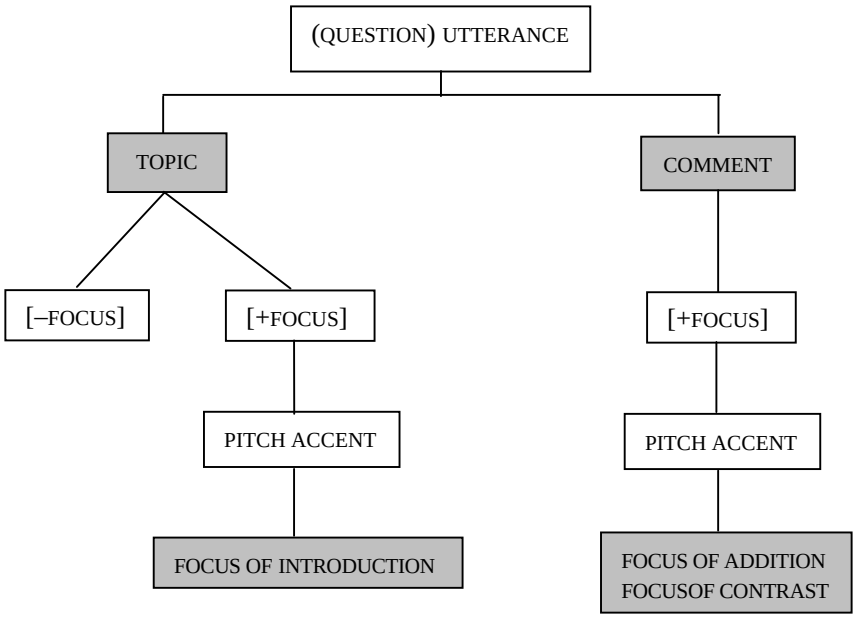


Figure 6.1. Focus domains ‘topic’ and ‘comment’, together with their focus types.

In chapter 5, we assumed that the relative scaling of successive accents gives expression to the underlying focus structure which, in turn, reflects the relative

significance of the corresponding constituents. We speculated that the biological Effort Code might be an important factor in the phonetic implementation of this focus structure. This universal code associates the size of an accent with the amount of vocal effort expended on it. By implication, increased effort is taken to reflect greater significance of the corresponding constituent, and vice versa (Gussenhoven 2000b). In the present chapter, we argued that statements and questions both consist of a topic and a comment, but that they differ with respect to the relative significance of these constituents. In statements, the communicative significance of topic and comment is broadly equal. In questions, by contrast, the comment is more significant than the topic, considering that it is the object of questioning which involves checking its correctness; hence it requires FOCUS OF CONTRAST. The acoustic correlates of this focus type differ from those of the other two focus types in that they enhance the salience of the focused constituent. In § 6.2.9, we look at these correlates in more detail, in an attempt to determine whether they may have brought about the asymmetric accent patterns observed in the questions. First, however, we want to ascertain whether an approach in terms of topic-comment structure may also be beneficial to analyses of accent distribution in statements.

6.2.8 Excursion: Topic-comment structure in statements

Analysing utterances in terms of topic/comment structure and allowing both of these to carry accents would seem to account for accent distributions in statements that have proved problematic. Consider a well-known minimal pair based on Halliday (1967a:38):

(6.5a) DOGS must be carried
(6.5b) Dogs must be CARRIED

The two different accent patterns are claimed to represent two possible readings of a written notice in the London underground. (6.5a) would seem to make it obligatory for passengers to bring dogs when taking the underground. By contrast, the assumption underlying (6.5b) is that *if* passengers have dogs with them, they are urged to carry their animals in view of the danger posed by trains and escalators. Gussenhoven (1984a:42) points out that, for the latter interpretation, (6.5b) needs an additional accent on *dogs* and that, in fact, three different accent patterns can be distinguished, each of which receives a different interpretation. Thus, (6.5b) can be split up into another minimal pair:

(6.6a) People will be SHOT
(6.6b) PEOPLE will be SHOT

either member of which might be uttered to prevent someone from trying out his new gun in the street. The assumption underlying (6.6a) is that, normally, there are people in the street, which makes it likely for some of these to be shot. In (6.6b), there is no such assumption. *People* is conditional here, that is, *if* there are people in the street, it may well be their fate to be shot. From the fact that, in the (b) sentence, Argument and

Predicate are both accented, Gussenhoven concludes that A and P have not merged into a single focus domain. That is, SAAR has failed to apply, presumably because of the conditional status of the Argument *people* (“ ‘If there is an A of this sort, then…’ ” cf. Gussenhoven 1984a:42). In order to account for the observed accent distribution, Gussenhoven appeals to the semantic feature [eventive], which marks sentences according to whether or not they refer to an event. He proposes to constrain the operation of SAAR such that it does not apply in [–eventive] sentences. This type of sentence can, in turn, be subdivided into a ‘contingency’ type (the above conditional sentence, cf. 6.6b) and a ‘definitional’ type (e.g. *MILK is ANIMAL*). Given the accent on the Predicate *animal* in the latter sentence, A and P have, likewise, failed to merge into a single focus domain.

In an analysis in terms of topic-comment (which permits two accent patterns), there would seem no need to appeal to the [–eventive] sentence categories ‘contingency’ and ‘definitional’. When the utterance is split up into a topic and a comment either of which may be [+focus], an accent may be assigned to both, as in (6.6b), or just to the comment, as in (6.6a). It must be added that when the topic has a conditional or a definitional status, an accent appears to be obligatory. This obligation may not seem to follow directly from the approach developed so far. Yet, considering that bringing a *topic* in focus indicates that, for some reason or other, this topic deserves the listener’s special attention, we would hardly expect a conditional or definitional topic to go *without* focus.

Ladd (1996:234) relates the different accent patterns in Halliday’s above pair (6.5a,b) to differences in phrasing. Thus, the one-accented version can be taken to constitute a single prosodic phrase, whereas the two-accented version has two phrases. As to why such differences in phrasing should at all exist, Ladd notes that “much of the reason for dividing sentences into two phrases or keeping them as one has to do with a general phenomenon of psychological or semantic ‘weight’”, but, he adds, “The sources of ‘weight’ are admittedly not very clear”. However, from a communicative perspective it would seem plausible for the pragmatic dichotomy topic-comment to represent just such a ‘source of weight’.

When the topic-comment approach allows utterances to be ‘topic-less’ (because the topic can be retrieved from the immediate context, cf. § 6.2.7 above), this may account for other ‘problematic’ accent distributions discussed in the literature (e.g. Gussenhoven 1984a, Ladd 1996), such as:

(6.7a) *JOHNSON died* vs. TRUMAN DIED (Schmerling 1976)

(6.7b) *The MILK is in the sun* vs. *MILK is ANIMAL* (Gussenhoven 1984a)

(6.7c) *Your EYES are red* vs. *Your eyes are BLUE* (Kraak 1970)

In order to account for the different accent distributions in the minimal pair in (6.7a), Schmerling labelled them as different sentence types, i.e. *JOHNSON died* as a ‘news sentence’, and *TRUMAN DIED* as a topic-comment sentence. Similarly, the accent differences in pair (6.7b) are analysed as reflecting a difference between sentences that are [+eventive] and [–eventive] (see above). As for the pair in (6.7c), Kraak

(1970; as discussed in Gussenhoven 1984a:32), considers both sentences as 'entirely comment'. However, given that the two accent patterns obviously differ, Kraak introduces a semantic feature [inherent property] of colour.

In the analysis outlined here, the first members of the three minimal pairs could be regarded as 'all-comment presentational utterances' (Gundel 1988). Thus, from *JOHNSON died* the overt topic has been deleted from the surface and the sentence constitutes a comment on the very broad topic of current affairs. Next, *The MILK is in the sun* is an all-comment sentence that may be spoken against the backdrop of a (sunny) domestic scene. Finally, *Your EYES are red* may be addressed as an all-comment sentence to a person who is apparently present and functions as a contextual topic. In all three cases, the topic can be retrieved and the comments have accents on their Arguments (cf. SAAR). By contrast, the second members of the above pairs have overt topics: *Truman*, *Milk*, and *Your eyes*, respectively. While the former two are [+focus], the latter is [–focus], which probably reflects that there is no need for the speaker to (re)introduce it to the discourse, being already part of it.

It would thus appear that an approach based on the universal topic-comment dichotomy is useful in the analysis of statements, too. For past analyses, the presence of one vs. two accents has raised some problems, and solutions involved an appeal to differences in sentence type, or an appeal to additional semantic features. However, problems would seem to be largely solved by a topic-comment analysis (together with the assumption that a topic needs not be overtly present), since this allows both one-accent and two-accent patterns. In point of fact, Ladd's suggestion that these two patterns reflect differences in phrasing might be extended to include the topic/comment dichotomy, with the latter serving as a universal basis for dividing utterances into (pragmatic) phrases.

6.2.9 Contrastiveness: Acoustic properties

In the literature, a good deal of attention has been paid to what has been variously called 'contrastiveness', 'contrastive prominence', 'contrastive stress' or '(contrastive) focus', and several phonetic correlates have been advanced for it[19]. Notably, contrastiveness is claimed to involve higher accent pitch, an intonational device that has also been associated with the signalling of emphasis[20] (e.g. Chafe 1976:35; Brown *et al.* 1980:29; Couper-Kuhlen 1984; Eady *et al.* 1986; Ladd & Morton 1997:322). In a series of experiments, Eady, Cooper and colleagues investigated the effect of contrastiveness on f0 patterns and duration. One of their main findings was that contrastive focus on a

[19] For an overview up till 1984, see Couper-Kuhlen 1984.

[20] However, there is no consensus in the literature as to whether a contrastive and an emphatic accent are identical. According to Brown *et al.* (1980:73), items can be heard as either contrastive or emphatic, suggesting two categories; yet, drawing a line between them proves difficult. In the view of Couper-Kuhlen (1984:157), contrastiveness must be seen as the prototypical form of semantic contrast; alongside it, emphasis is less prototypical because it does not express a strict contrast in the Chafian 'instead of' sense (Chafe 1976). Thus, the 'contrastive' relation expressed by emphasis is felt to be looser and less specific. Grønnum (1990:32), on the other hand, speaks about 'emphasis for contrast', implying that emphasis is a means for expressing contrast. It would seem that Chafe (1976:36) is of roughly the same opinion when he says: "Probably this increased prominence often given to a contrastive focus results from an increased emotional commitment that is likely to accompany contrastiveness".

lexical item results in a raised f0 peak and an increase in duration for the contrastively focused word. These acoustical attributes proved fairly constant, regardless of whether they occurred in one or in more items in the same utterance (e.g. Eady & Cooper 1986, Eady *et al.* 1986).

That speakers clearly encode contrastiveness and that listeners are well able to identify a contrastively focused item was found by Brown *et al*. (1980:152). Speakers produced statements and questions with neutral accents (read speech) and with contrastive accents (semi-spontaneous speech). The latter were elicited by having participants reconstruct a story by asking yes-no questions based on a list of characters and actions. However, this list was jumbled on the idea that this would force the interlocutors to contrast the given items with one another. Measurements showed that, in utterances featuring such contrastive items, the three phonetic cues 'maximum pitch height', 'maximum pitch movement' and 'maximum intensity' mostly cumulated in the contrastive item. In the neutral utterances, on the other hand, these properties tended to spread over several different items. Crucially, these production differences corresponded with differences in perception. When listeners were asked to mark the 'tonic(s)'[21] in both types of utterances, they readily singled out the contrastive item, even from quite long structures. By contrast, in the neutral utterances judgements were more or less erratic. Even in the short sentences, at least two but mostly as many as three to five words were marked as tonics; in fact, as the authors note, any item perceived as stressed seemed 'at risk' (p. 146). These results indicated that the contrastive accents were phonetically distinct from the non-contrastive accents, in production as well as perception.

A non-local property of contrastive intonation is the reduction of accompanying accents that are not contrastively focused. According to Grønnum (1990:32), "[…] highlighting is a relative phenomenon and may be obtained through shrinking/lowering/reducing surrounding f0 events […]", and "[…] the minimal requirement for perceiving a contrast is the reduction of surrounding stress groups […]". This amounts to saying that, in multi-accented utterances, contrastiveness produces asymmetric accent patterns. Contrastiveness has further been associated with a steep fall from a relatively high pitch level (e.g. Chafe 1976:36; Couper-Kuhlen 1984). Particularly in sentence-initial contrasts, this fall was found to be more extensive than the preceding rise (Eady *et al.* 1986:248). For the purpose of the present study, two perception experiments seem of particular interest, because they involved two-accented utterances with contrastive focus. Therefore, we report them in greater detail.

The experiment carried out by Rump & Collier (1996)[22] was based on the two-accented Dutch statement *aMANda gaat naar MALta* ('Amanda goes to Malta'). Listeners had to adjust the height of either accent peak relative to the other (which was fixed) so as to make the utterance match with one of the following focus structures: (i) neutral focus (answering the question 'What is happening?'), (ii) contrastive focus on 'Amanda' (answering the question: 'Is John going to Malta?'), (iii) contrastive focus on 'Malta' (answering the question: 'Is Amanda going to Cyprus?'), and (iv) double focus (answering the question: 'Is John going to

[21] I.e. the nucleus, or main accent. All subjects taking part in the experiments had declared to be familiar with the concept of 'tonic'.

[22] This experiment was discussed earlier in chapter 4 (§ 4.3.2.1).

Cyprus?'). Adjustments were made in the following way. While the fixed peak had one out of three possible f0 values, subjects had to choose the appropriate f0 value for the adjustable peak from a peak height continuum divided into nine steps (peak 2) or eleven steps (peak 1) of 0.25 ERB (about 1.5 semitones). Across the four focus conditions, the declination rate was kept constant. On the basis of the adjustments, prototypical accent patterns of each focus structure were set up; these are shown in Figure 6.2.

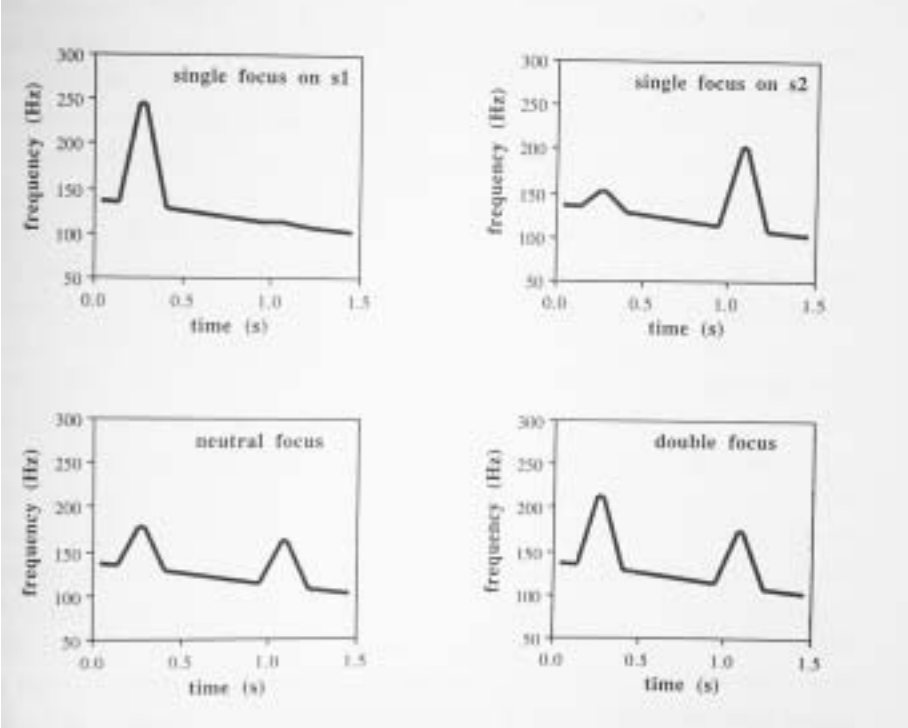


Figure 6.2. Prototypical accent patterns in four focus conditions, constructed on the basis of adjustments (see text). Figure taken from Rump & Collier 1996.

Next, (new) subjects listened to 48 versions of the same test sentence in which the peak heights of the two accents had been systematically varied. For each version, listeners had to decide to which of the four focus-eliciting questions it was the likeliest answer. The results showed that the subjects had clear ideas about the acoustic characteristics of each focus structure. Contours with roughly equal peaks were judged to have 'neutral focus' (as long as both peaks were relatively low). When both peaks were roughly equal but had relatively high peaks, this was identified as contrastive focus on both items ('double focus'). Asymmetric patterns with a step-up or step-down in pitch were perceived as having (single) contrastive

focus on the higher-peaked accent. Importantly, in the latter two focus structures one accent had to be substantially reduced. Thus, when the contrastive accent occurred in *final* position it had to be preceded, ideally, by an accent that was considerably smaller than its counterpart in neutral focus. In *initial* position, the contrastive accent was even followed by a stretch of almost level pitch: the accentable syllable of the second accent virtually lacked any pitch movement (an effect also referred to by many others, e.g. Eady and Cooper 1986). The results of this second experiment validated the prototypical contours set up on the basis of the peak adjustments in the first experiment. In both experiments, the subjects easily kept the four focus conditions apart on the basis of pitch characteristics; crucially, asymmetric accent configurations signalled contrastiveness on the high-pitched item.

An earlier perception experiment by Bartels & Kingston (1994) provided similar evidence that asymmetry between successive accents forces an interpretation of contrastiveness on the high-pitched item. The experiment was based on Pierrehumbert & Hirschberg's (1990) claim that L+H* conveys contrastiveness, whereas H* is merely neutrally assertive. The bitonal L+H* pitch accent was assumed to specify a low target followed by a steep and rapid rise. By contrast, the monotonal H* rose less steeply from a mid level. Bartels & Kingston wanted to establish how listeners keep these two pitch accents apart. That is, if L+H* indeed conveys contrastiveness, which acoustic properties contribute most to a contrastive interpretation: (i) the presence of a low-pitched L tone (i.e., an unambiguous L+H* pitch accent), (ii) the peak f0, (iii) the temporal alignment of the L tone (if present), or (iv) the temporal alignment of the peak with the stressed syllable? In the two-accent test utterance *aMANda had a baNAna, amanda* always featured H*, whereas the accent on *banana* was systematically varied with respect to the above four variables. This meant that the accent was either H* or L+H*, depending on the presence of a low target. Subjects were given a certain scenario. Upon hearing the manipulated versions of the test utterance, they had to decide whether or not these implied contrastiveness (first part), and whether they implied broad or narrow focus (second part).

Our chief interest, however, is in the extensive pilot studies carried out prior to the main experiment. Here, Bartels & Kingston adopted the prototypical peak heights of L+H* and H* as gleaned from several pitch track corpora. This meant that the peak of L+H* was, on average, 60-100 Hz. higher than the peak of H*. As a result, the test utterance *aMANda had a baNAna* displayed a substantial step-up in pitch between the first and the second accent. This higher peak on baNAna completely obliterated the effects of the other phonetic features that had been manipulated. Bartels & Kingston also found that the cross-over point, in terms of peak height, from the 'non-contrastive' to the 'contrastive' interpretations was impressively stable. Given this overwhelming effect of peak height, they decided to restrict peak height variation in the main experiment to the very narrow window of 16 Hz. Still, they were aware that any effect of the three other manipulations was likely to be only secondary to the observed dominant effect of peak height. This expectation proved correct: of the four variables, peak height remained the strongest cue to contrastiveness. That is, the experiment did not truly confirm that a contrastive interpretation is forced by the presence of an L target whereas the monotonal H* cues a non-contrastive interpretation. Neither did peak alignment or rise onset alignment have systematic effects on the perception of

contrastiveness. Instead, the difference between contrastive and non-contrastive interpretation seemed to be chiefly signalled by the greater peak height of the contrastive item.

The perception experiments of Rump & Collier and Bartels & Kingston are relevant to our data since they demonstrate that, at least in two-accented statements, contrastiveness is typically bound up with asymmetric accent patterns. While the contrastively focused accent is raised, the other accent may be reduced or altogether absent. In an initial accent, contrastiveness frequently shows up as a steep fall from a relatively high f0; it is typically followed by an almost level stretch of pitch. While in statements contrastiveness must be seen as a marked option, we propose that this is not so in questions. Earlier in this chapter, we argued that the comment in questions represents the speaker's presupposition which is the object of questioning. This comment is put forward with a view to obtaining the correct item from a set of potential comments. Because of this, the focus type qualifies as FOCUS OF CONTRAST. Accordingly, in questions contrastiveness of the comment can be regarded the unmarked situation.

If this proposal is generally on the right track, we expect the various accent patterns in our data to have differed as a function of utterance type and, by implication, of focus structure and the corresponding differences in phonetic correlates. Basing ourselves on the approach developed so far, we will establish focus domain(s) and focus type(s) for each of the core utterances and predict the corresponding accent patterns. The predicted patterns are then compared with the actual patterns as they were realised by the speakers.

6.3 Focus and accent in the experimental utterances

6.3.1 Statements

Focus structure

In the statements *Renée heeft nog vlees over* ('Renée has some meat left') and *Marina wil haar mandoline verkopen* ('Marina wants her mandolin sell'), the referents of the proper names function as topics; they are supposed to be uniquely identifiable objects in the interlocutors' common background. Nonetheless, in so-called 'all-new' statements it may be necessary for these topics to be first established in the listener's current awareness, before any comment pertaining to them can be added ('I'm going to bring up something respecting *Marina/Renée*'). In that case, the topics are [+focus] (i.e. FOCUS OF INTRODUCTION) and receive an accent. Next, the respective comments (*heeft nog vlees over* and *wil haar mandoline verkopen*) add new information to the (previously established) topics. This means that they are, also, [+focus] (i.e. FOCUS OF ADDITION) and carry accents. SAAR then predicts that the Arguments *vlees* and *mandoline* are going to be accented.

Expected intonation contour

When the statement is uttered without a particular context, the discoursal weight of the two focus domains is more or less equivalent: 'I want to bring up something about Marina (topic), viz. that she wants to sell her mandolin' (comment). Since there seem to

be no indications that the two focus types FOCUS OF INTRODUCTION and FOCUS OF ADDITION are associated with specific phonetic properties, we expect the two accents in the statements to be roughly similar (cf. Rump & Collier's prototypical neutral statement contour in Figure 6.2).

Observed intonation pattern(s)
In chapter 4, it was seen that the difference between the excursion sizes of the consecutive accents was smallest in the statements (Figure 4.13). Also, the statements had a relatively slight inter-peak interval[23], viz. a mean of –0.01 ERB in the men and a mean of 0.21 ERB in the women. Instantiations of this recurring pattern are given in Figure 6.3.

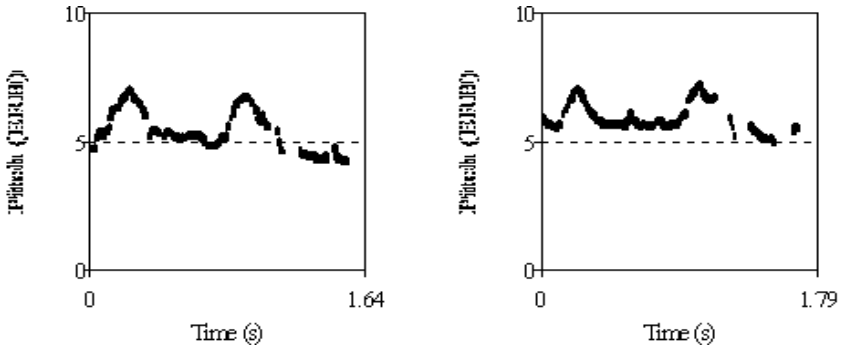


Figure 6.3. Statements: Default accent pattern of *maRIna wil haar mandoLIne verkopen* (%L H*L H*L L%) spoken by ES and DB (female).

6.3.2 Wh-questions

Focus structure
In the wh-question *Waar wil Marina haar mandoline verkopen?* ('Where wants Marina her mandolin sell?'), the subsentence 'Marina wil haar mandoline verkopen' is already part of the speaker's background. Hence, it constitutes the topic of the utterance ('*As for* Marina wanting to sell her mandolin'). As noted earlier, a topic does not necessarily coincide with the subject or even with a noun phrase. In the present case, it comprises a full sentence. To this, the questioner adds the comment *Where?*, reflecting his presupposition that Marina wants to sell her mandolin *at some place*. The wh-phrase further indicates that the speaker wants to extend his background knowledge to *include* this place, i.e., he wants to be provided with the correct item from a set of possible locations. Hence, the appropriate focus type is FOCUS OF CONTRAST.

[23] To be sure, in WQ this interval was even smaller, but in this question type subject and object accent were part of the presupposed topic sentence, see § 6.3.2 below.

Expected intonation contours

As argued in § 6.2.1, the topic (here: a full-blown sentence) may or may not be in focus, depending on whether the speaker assumes it to be already part of the listener's current awareness. When it is [–focus], we expect its subject and object to be without accents; this would be option 1. Considering that the initial wh-word has FOCUS OF CONTRAST, this is predicted to have a relatively large accent which is not followed by other accent peaks (cf. panel 'single focus on s1' in Figure 6.2 above). Hence, we expect the strongly left-asymmetric accent pattern as illustrated in Figure 6.4.

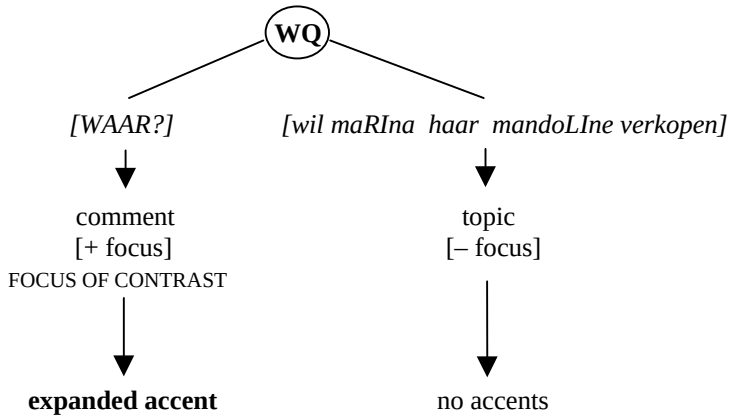


Figure 6.4. Focus structure and predicted accent pattern in WQ, option 1.

In the other option, i.e. when the topic sentence needs to be (re)introduced to the discourse and is therefore [+focus], its focus type is obviously FOCUS OF INTRODUCTION. As this topic comprises a complete statement with a topic and a comment of its own, it may have either of the two accent patterns available to statements (see above). Again, we expect the initial wh-word (which functions as the comment and is in FOCUS OF CONTRAST) to feature an expanded accent. In addition, accents in the topic sentence are likely to be reduced, given that contrastiveness is typically enhanced by reduction of neighbouring accents. Consequently, the expected accent pattern in option 2 is also left-asymmetric, be it to a lesser extent than in option 1. Figure 6.5 illustrates this.

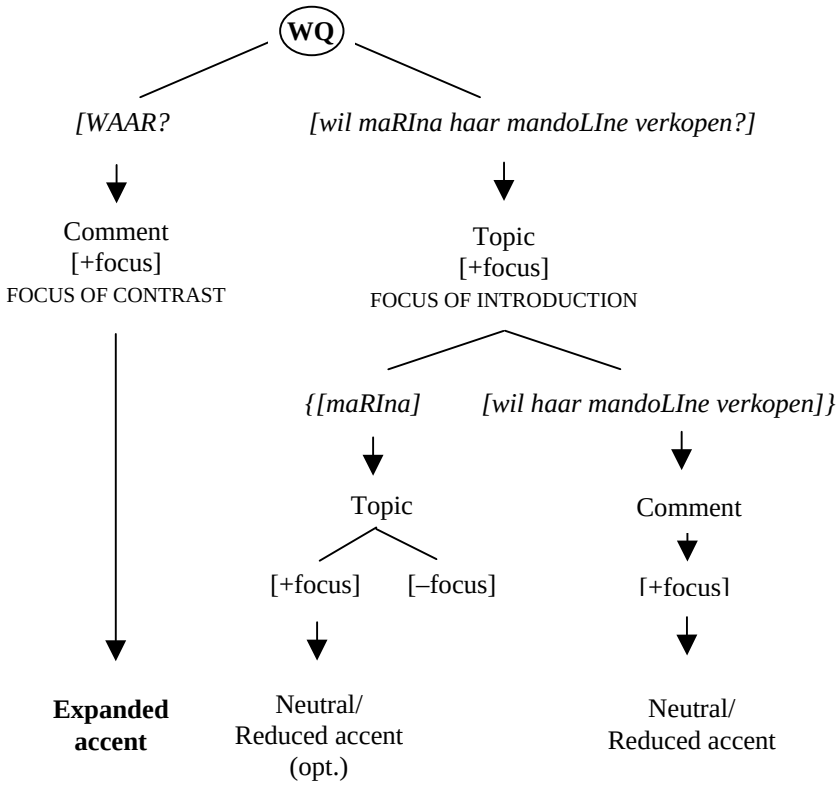


Figure 6.5. Focus structure and corresponding accent pattern in WQ: option 2.

By analogy, the above line of reasoning also applies to the other wh-question in the corpus, *Wat heeft Renée nog voor vlees over?*. Strictly speaking, in this question the wh-word is part of the discontinuous phrase *Wat […] voor vlees?* ('what type of meat?'). However, moving the entire phrase (which functions as the direct object) into initial position does not bring about differences in meaning or accent location.

Observed patterns

Together, the left-asymmetric patterns predicted by the Figures 6.4 and 6.5 occurred in 71% of the 186 neutral tokens of WQ (see chapter 4, Figure 4.3). As reported earlier, subjects and objects were often without accents (in 88% and 37%, respectively), which limited the occurrence of option 2. Clearly, most speakers

preferred to make the comment (the question word) relatively prominent alongside (reduced or non-existent) accents in the topic sentence. In fact, the f0 pattern of option 1 (no accents on subject and object) closely resembled the pattern of *statements* with initial contrastive focus in Figure 6.2. Rump & Collier (1996) attribute this deaccenting after initial contrastive focus to an effort on the part of the speaker to de-emphasise this portion of the sentence. In option 2 (Figure 6.5), the relatively large wh-accent was followed by one (or two) smaller accents in the subsentence. Following Bolinger (1972a:137), we could look upon this pattern as an instance of intonational embedding, such that "the whole intonational scale is reduced for a given stretch of speech, forcing that stretch to be interpreted in contrast to the whole utterance". That is, within the reduced stretch there are normal syllable-by-syllable contrasts, but in the question as a whole, the reduced stretch is obviously of less significance.

In the contours in Figure 6.6, the difference between the focus structures of option 1 and option 2 stands out clearly.

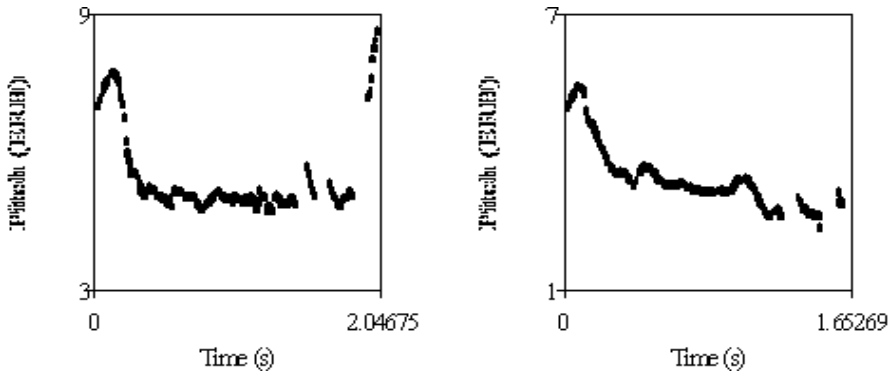


Figure 6.6. Left-asymmetric accent patterns in the wh-question *WAAR wil maRIna haar mandoLIne verkopen?*.
Lefthand panel: option 1 (%L H*L H%), speaker FL (female).
Righthand panel: option 2 (%L H*L !H*L H*L L%), speaker MC (male).

In all, the asymmetric patterns of options 1 and 2 displayed the phonetic characteristics typical of contrastiveness, suggesting that the comment (the wh-phrase) was in FOCUS OF CONTRAST. While the peak of the wh-phrase was relatively high and its fall mostly steep and extensive, there was also a substantial reduction of the neighbouring accents in the topic sentence. This set of features proved effective in bringing out a perceptual contrast between the topic and comment constituents, witness the clear impression of auditory subordination reported in chapter 4.

However, we also observed a *non*-predicted pattern, i.e. the one exemplified by Figure 4.4 (chapter 4). Here, the H* of the wh-phrase formed a high plateau with the peak of the object accent (i.e. H* H*L). This configuration occurred in 29% of the wh-questions; in roughly half of these cases, pitch made a further step-up from

the high plateau to mark the object (i.e. H* ^H*L). In this pattern (which tended to be rather speaker-specific), the topic sentence was not prosodically subordinated to the comment. Nonetheless, it featured a high pitch plateau. In chapter 5, such plateaus were taken to be motivated by the Frequency Code and interpreted as underscoring the question status of an utterance. Pragmatically, the interpretation of the plateau pattern might slightly differ from that of the patterns in Figure 6.6. In the latter two, the topic sentence is either fully taken for granted, or it is reintroduced to the discourse. The questions can be paraphrased as 'As we both know that Marina wants to sell her mandolin somewhere, can you provide me with the correct location?', and 'I suppose, but I'm not fully certain, that we both know that Marina wants to sell her mandolin somewhere, but either way, can you provide me with the correct location?', respectively. In the plateau pattern, however, the topic-sentence might have a conditional status (cf. the discussion in § 6.2.8 above), that is, it could be paraphrased as 'If it is true that Marina wants to sell her mandolin somewhere, can you provide me with the correct location?'. In this reading, an accent on the topic-sentence is obligatory. Figure 6.7 gives an example of the high-plateau contour.

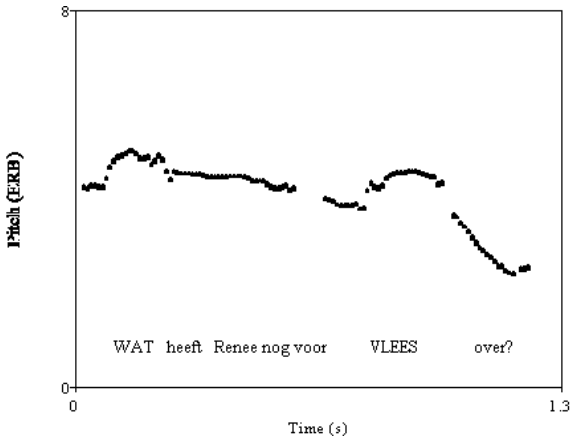


Figure 6.7. Accent pattern of the wh-question *WAT heeft Renée nog voor vlees over?* featuring a high pitch plateau between wh-phrase and object accent (%L H* H*L %L); speaker HD.

6.3.3 Yes-no questions

Focus structure

For the yes-no questions in the corpus, the argumention is essentially the same. In *Wil Marina haar mandoline verkopen?*, *Marina* constitutes the topic ('As for *Marina,…*'). With respect to this topic, some comment is being questioned. This comment represents the presupposition currently entertained by the questioner, which may be understood as 'At present, I believe the referent of the topic wants to sell her mandolin'. It is thus put forward as the correct comment out of a set of other potential comments pertaining to Marina. As the motive underlying focus is to establish the correctness of the

presupposed comment, the focus type is FOCUS OF CONTRAST. According to SAAR, the comment's Argument (i.e. *mandoline*) receives the accent. As regards the topic, when the speaker has reason to suppose that the referent of *Marina* is not currently part of the listener's awareness, the topic can be (re)activated by making it [+focus] (FOCUS OF INTRODUCTION). When, on the other hand, it is clear from discourse or context that *Marina* is already present in the listener's mind, no accent on the topic is needed. The same applies to the other experimental yes-no question, *Heeft Renée nog wat vlees over?*.

Expected intonation contour

When topic and comment are both [+focus] and accented, the intrinsic acoustic difference between the two focus types (FOCUS OF INTRODUCTION vs. FOCUS OF CONTRAST) is expected to cause pitch to step up between topic and comment. The step size may be further increased by a reduction of the subject accent (cf. panel 'single focus on s2' in Figure 6.2 above). Hence, the accent pattern is expected to display right-asymmetry, as shown in the diagram in Figure 6.8.

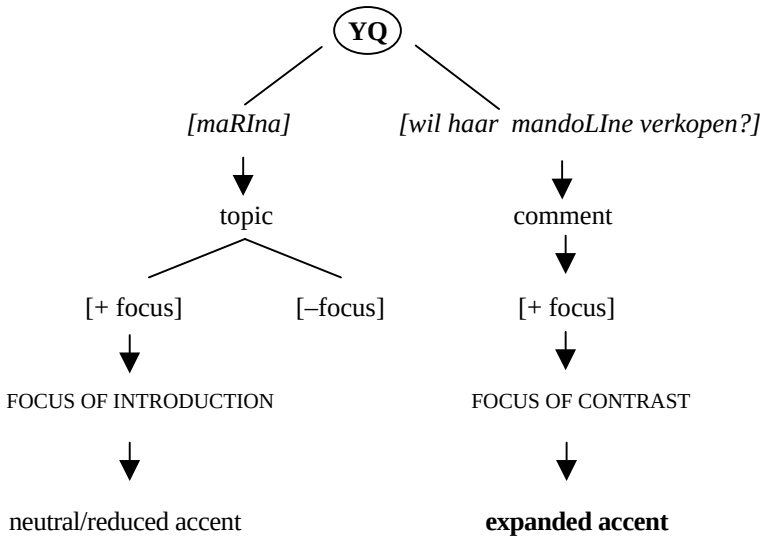


Figure 6.8. Focus structure and predicted accent pattern in YQ.

Observed patterns

YQ showed the predicted right-asymmetry in 59% of the 199 tokens. Note that, in 29% of the tokens, occurrence of the predicted pattern was precluded because the pitch accent on the object was L*(H) or because the subject was unaccented. In effect, in only 12% of the tokens with the sequence H*(L) H*L did the expected asymmetric pattern

fail to occur. The average amount of upstep was 0.43 ERB in the men, 0.88 ERB in the women; the difference with statements was significant (cf. chapter 4, § 4.3.3.2). Figure 6.8 shows pitch contours of the unlinked and the linked pattern, both of which feature the predicted upstep in pitch.

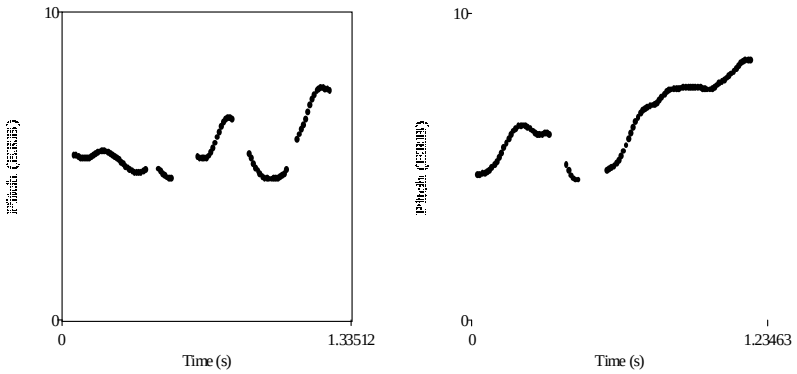


Figure 6.9. Yes-no questions: Right-asymmetric patterns in *wil maRIna haar mandoLIne verkopen?*, speaker ES (female).
Lefthand panel: L% H*L ^H*L H%. Righthand panel: L% H*L ^H H%.

6.3.4 Declarative questions

Focus structure and expected intonation pattern
Although the declarative questions (*Marina wil haar mandoline verkopen?* and *Renée heeft nog vlees over?*) differed from the YQ versions in that they lacked inversion, we assumed earlier that the two question types have the same focus structure (cf. Figure 6.7). However, it is precisely because of this absence of inversion that DQ fully relies on intonation for the expression of questionhood. Accordingly, we expect greater expansion of the object accent and stronger reduction of the subject accent, since this would further increase the size of the upstep between topic and comment.

Observed patterns
In DQ, the predicted upstepping pattern occurred in 76% of the 198 tokens. In line with the predictions, in both sexes the mean inter-peak intervals increased as a function of question type (Table 6.1); the differences were not significant, though (cf. chapter 4, Table 4.18).

Table 6.1. Mean intervals (ERB) between accent peaks on subject and object (upstep) in yes-no questions (YQ) and declarative questions (DQ), broken down by Sex of speaker.

Question type	Male	Female
YQ	0.43	0.88
DQ	0.61	1.12

Moreover, the asymmetric pattern occurred more frequently in DQ than in YQ (76% vs. 59%, respectively). However, DQ did not meet the expectations on all counts. The mean excursion on the subject accents was somewhat less reduced than in YQ, rather than more, as had been expected (cf. chapter 4, Figure 4.9). Also, full deaccentuation of the subject occurred only in six DQ tokens, against 32 in YQ. Apparently, in DQ the discoursal weight of the topic was slightly greater than in YQ. In chapter 5, the results of a pen-and-paper experiment made it clear that DQ may be used for two purposes, viz. to gain information and to gain confirmation. Arguably, a confirmation question wants confirmation of the ensemble of comment *and* topic, whereas an information question primarily seeks the correctness of the comment respecting the topic. The dual function of DQ may account for the slight acoustic differences between the two question types which were otherwise largely similar. Figure 6.10 gives two sample contours of DQ. In the lefthand panel, the object accent H*L is fully realised; in the righthand panel, its L tone is deleted and its peak is directly linked to the high final boundary tone H% (cf. the discussion on high plateaus in chapter 5, § 5.3.2.2).

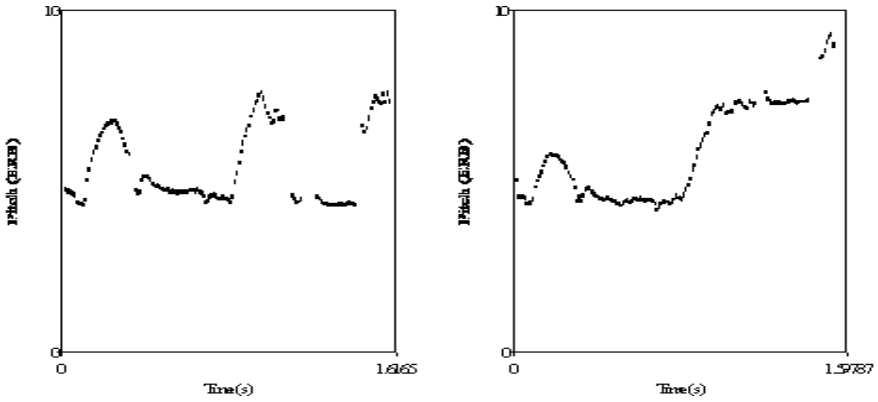


Figure 6.10. Declarative questions: Right-asymmetric patterns in *maRIna wil haar mandoLIne verkopen?* (speaker ES, female)
Lefthand panel: %L H*L ^H*L H%. Righthand panel b: %L H*L ^H* H%.

6.3.5 Discussion

6.3.5.1 Accent asymmetry and focus type

Overall, the accent patterns produced by the ten speakers matched the patterns predicted on the basis of focus structure and focus type. While, underlyingly, the contours usually consisted of occurrences of the phonological category H*L, the phonetic implementation of this category displayed substantial variation across utterance types. This variation may be attributed to differences between statements and questions as regards their focus types. In all statements, both the topic and the comment were accented and hence [+focus]. Since the corresponding focus types (i.e. FOCUS OF INTRODUCTION and FOCUS OF ADDITION, respectively) have similar acoustic correlates, the statements typically featured the broadly symmetric accent patterns shown in Figure 6.3. By contrast, in the questions a majority of the accent patterns were overtly asymmetric, due to the acoustic characteristics of the focus type of the comment, FOCUS OF CONTRAST. Thus, in YQ and DQ there was right-asymmetry, produced jointly by a relatively small subject accent (representing the topic) and a relatively large object accent (representing the comment). Conversely, WQ mostly showed left-asymmetry, caused by the pitch drop between the relatively large initial wh-word (representing the comment) and the topic sentence, which was either unaccented or had reduced accent(s). In a minor pattern, there was no asymmetry due to a high plateau between the peaks of wh-word and object accent. Table 6.2 recapitulates the results.

Table 6.2. Percentages of accent asymmetry, together with the corresponding intervals (ERB) between subject peak and object peak, broken down by utterance type.

	Asymmetric accent pattern in	Corresponding inter-peak interval (ERB)
Statements	27%	+.10
Wh-questions	71%	– .28
Yes-no questions	59%[24]	+ .66
Declarative Questions	76%	+ .87

The match between predicted and observed patterns would seem to support the proposed approach to focus in questions, which claims that the appropriate focus type for the comment constituent is FOCUS OF CONTRAST. This focus type is typically accompanied by expansion of the relevant accent and reduction of neighbouring accents. Crucially, our approach only predicted asymmetry of the accent patterns in questions, it did not specify the actual direction of the bias. That the three experimental question types displayed two opposing patterns, a right-asymmetric and a left-asymmetric, reflects that this bias is simply a function of the position of the comment within the question utterance.

[24] As noted earlier, in 29% of the YQ tokens upstep between the accents (and hence accent asymmetry) was precluded, either because the subject was unaccented or because object accents had an L*(H) pitch accent rather than H*(L) (cf. § 6.3.3).

6.3.5.2 Focus type and the Effort Code

In chapter 5, two out of the three manifestations of high(er) question pitch ('Q=H') were taken to derive from the universal Frequency Code. The third however, i.e. the asymmetric accent patterns, could not be explained on the basis of the FC. The reason was that, across the three question types, tangents drawn through the accent peaks showed opposing directions: upward in YQ and DQ (compatible with the FC), but downward in WQ (incompatible with the FC).

In *statements*, sequences of differently scaled accents typically reflect a **non**-neutral focus structure (e.g. Liberman & Pierrehumbert 1984, Bartels & Kingston 1994, Rump & Collier 1997). As suggested in chapter 5, focus structure can, ultimately, be regarded as a grammaticalisation of the biological Effort Code. This universal code embodies a universal form-function relation between the local range of an accent and the relative significance of the corresponding constituent. Speakers may express this gradient meaning in the phonetic implementation of phonological categories (Gussenhoven 1999b, 2002 to appear). As a result, two consecutive H*L pitch accents may display asymmetric patterns, such as those exemplified in Figure 6.2 above.

In chapter 5, we were not yet in a position to determine whether the asymmetric accent patterns in the questions could be attributed to the workings of the Effort Code. First, we had to delve more deeply into the issue of focus in, particularly, questions. In the present chapter, however, we proposed an account of focus which brings the three question types together in a second generalisation[25]. Assuming that WQ, YQ and DQ involve requests for the correct item from some (implicit or explicit) set, we proposed that questions differ from statements in the distribution of focus type. Given their specific discoursal function, questions require the focus type FOCUS OF CONTRAST, whose phonetic correlates then determine the overtly asymmetric patterns of accentuation. That scaling differences between successive accents are meaningful is, obviously, precisely what the Effort Code would predict. Accordingly, the two different asymmetric accent patterns in our question material can be captured by a single account based on the Effort Code. This means that the conclusion of chapter 5 can now be extended to include the EC: question intonation crucially involves the operation of *two* biological codes, the Frequency Code and the Effort Code, whose gradient meanings are (largely) expressed in the phonetic implementation.

As noted earlier, pitch modifications on the basis of the Frequency Code and Effort Code mainly affect pitch range (cf. chapter 5, § 5.3.5). While the FC would seem to modify, broadly, the overall *level* of pitch (final rise, high plateaus, raised register level), the EC clearly involves pitch *span* (e.g. focus structure/emphasis). The codes also seem to be associated with different *types* of gradient meaning, which may explain why their pragmatic effects would not appear equally forceful. Thus the global variation in register level motivated by the FC seems less finely-tuned to meaning differences than the local variation in accent range motivated by the EC, which may express differences in focus structure. On the

[25] The first generalisation stated that, relative to the statements, the questions had a greater overall share of high pitch (i.e. Q=H).

whole, the gradient differences in register *level* would seem to have a mainly supporting role, in that they may render the intonation of a given utterance more prototypical.

6.4 Summary and conclusions

Focus structure and the corresponding accent pattern help the listener interpret the discoursal content of a message; they turn a sentence (a semantic/syntactic notion) into an utterance (a pragmatic notion). The object of the present chapter was to gain insight into the focus structures underlying the accent patterns in our production corpus. Considering that these patterns differed as a function of the opposition statement-question, our first research question asked whether the function of focus in questions is different from that in statements. This led to the following account.

At the pragmatic level, utterances are to be analysed in terms of the universal dichotomy topic vs. comment. Either of these complementary constituents constitutes a domain for [+focus] or [–focus]. Since speakers may have various reasons for bringing items in focus (which may have phonetic consequences for accent scaling), we distinguished three focus types:

(6.8)

i. FOCUS OF INTRODUCTION, i.e. the focussed part has to be (re-)introduced to the discourse, the speaker's current awareness, etc.
ii. FOCUS OF ADDITION, i.e. the focussed part adds information to the topic;
iii. FOCUS OF CONTRAST, i.e. the focussed part is the 'correct' member of an (implicit or explicit) set (cf. Chafe 1976);

In statements as well as in questions, the *topic* reflects 'what the utterance is about'; [+focus] status and accent are optional, depending on whether or not the topic has to be explicitly (re)introduced to the discourse (FOCUS OF INTRODUCTION). As for the *comment*, in statements a speaker may bring this in focus (i) to add something to the topic (FOCUS OF ADDITION), or (ii) to contrast the comment with other members of a set of potential comments (FOCUS OF CONTRAST). Questions, however, would seem to have use only for the latter type of focus: here, the comment does not mean to *add* anything to the topic. Instead, in yes-no questions and declarative questions (YQ, DQ), the speaker checks whether the comment proffered by him is the correct one out of an implicit set of potentially appropriate comments (one of which is the negation of the uttered comment). Thus, in the yes-no question *Wil Marina haar mandoline verkopen?*, the speaker gives expression to a comment that is currently entertained by him with respect to the topic 'Marina', viz. that the referent of this topic wants to sell her mandolin. Being aware that this comment may be (partially) incorrect, he is intent on being provided with the correct comment. In wh-questions (WQ) the comment, embodied by the wh-phrase, offers the listener a naturally constrained set of choices, one of which is expected to be correct. Thus, in the wh-question *Waar wil Marina haar mandoline verkopen?*, the comment 'Where?' indicates that the topic sentence 'Marina

wil haar mandoline verkopen' can be extended to include some location. In statements as well as in questions, the comment requires a [+focus] status and, hence, an accent. In sum, comments in questions typically differ from comments in statements in the way focus *functions*. Consequently, they also differ as regards the distribution of focus *type*.

Our second research question asked whether a difference between statements and questions in the function of focus might have produced the observed differences in accent pattern. In the account outlined here, the unmarked focus type in question comments is FOCUS OF CONTRAST. The literature associates contrastively focused items with, notably, expansion of pitch movement and reduction of neighbouring accents. These phonetic properties were amply observed in the experimental questions, not in the statements. They were taken to be motivated by the biological Effort Code, which associates larger pitch movements with greater significance, and vice versa, and they adequately accounted for the opposing asymmetric accent patterns found in the three question types.

The assumption that a topic may be [+focus] as well as [–focus] allows questions and statements to have two basic accent patterns which, between them, reflect certain pragmatic differences. It also accounts for a number of problematic accent patterns in *statements* that have been discussed in the literature.

The analysis developed here crucially rests on the universal topic-comment dichotomy. Considering that questioning is pre-eminently a discoursal category, building one's analysis on these pragmatic constituents would only seem adequate. It must be emphasised, however, that the proposed approach to focus in questions is no more than a first step. To arrive at a clearer understanding of this specific subject matter, many further steps will have to follow.

Chapter 7

Summary and conclusions

7.1 Introduction

So far, work on standard Dutch intonation has largely concentrated on the declarative utterance, presumably because this can be regarded as the default utterance type. The present study focused on a discourse category which is, also, of considerable communicative importance, the interrogative utterance. The study compared systematically varied production data of Dutch statements and their question versions (covering three different question types) with a view to identifying the latter's intonational properties. That is, throughout the study it was assumed that the intonation of statements represents the default situation from which question intonation may deviate. The experimental corpus comprised 200 statements (ST), 200 wh-questions (WQ), 200 yes-no questions (YQ) and 200 declarative questions (DQ), read out by five female and five male speakers. Given the exploratory character of the research, hypotheses were largely formulated on the basis of studies of question intonation in languages other than Dutch. These hypotheses were tested and the Dutch question data were found to share a number of features with comparable data in other languages. The main characteristic of these features was that, across the board, pitch in the questions was higher than in the corresponding statements. This consistent association between question intonation and high(er) pitch suggests a greater degree of iconicity than is usually allowed for in linguistic theory. Furthermore, comparison of the accent patterns in the statements and questions led to a discussion of focus in questions and, more specifically, of the interaction between focus and question intonation. So far, this issue has not attracted much attention in intonation theory.

In the present chapter, the main observations and findings are recapitulated, first against the background of the central research questions laid down in chapter 1 (§ 7.2.1), next from the perspective of the hypotheses formulated in chapter 2 (§ 7.2.2-6). In § 7.2.3, we move on to additional observations and considerations to which the data gave rise. Section 7.3 discusses some theoretical and practical implications of the results, and in § 7.4, suggestions are made for further research in the area of Dutch question intonation.

7.2 Main observations and findings

7.2.1 Research questions

The present study concerned itself with the following research questions:

- Does the discoursal function 'interrogativity' systematically correspond with specific intonational properties?
- If this is the case, are these properties phonological, phonetic, or both?

The first question could be straightforwardly answered in the affirmative: the intonational expression of the discoursal function interrogativity typically involved the feature 'high(er) question pitch'. To give an example, in all three question types, f0 values of phonologically low tones lay well above the corresponding values in the statements. This overall finding, summed up by the formula Q=H, was in agreement with a wide range of question data from related and unrelated languages, suggesting that it was not accidental. Rather, it seemed plausible for Q=H to have derived from the universal, innate Frequency Code (Ohala 1983, 1984, 1994). This biological code associates high pitch in a vocaliser with small body size and, by extension, with submissiveness and dependence, whereas low pitch is associated with large body size and, by extension, with authority and assertiveness. Although, in linguistic theory, appeals to iconicity and innateness are not always looked upon with favour, this reservation would seem to apply less to intonation, where the dividing line between the linguistic system and the paralinguistic or extralinguistic systems is less rigid. As Gussenhoven (2002 to appear) puts it, "the Frequency Code […] will shine through the chinks of the linguistic system".

The second research question asked whether the Q=H effects manifested themselves in the phonology, in the phonetics, or in both. On the whole, interrogativity did not appear to affect the phonological system in that questions require specific pitch accents and/or boundary tones. Thus, across the experimental material, i.e. in the statements as well as in the questions, the predominant (surface) pitch accent was H*L (statements: 87%, questions: 72%), and the predominant *initial* boundary tone was %L (statements: 96%, questions: 89%). To be sure, the questions typically featured the high *final* boundary tone H% (87%) whereas the statements did not (0%). Although this latter distribution was no doubt systematic, it is also obvious that utterance-final high f0 is not exclusively tied to interrogativity. Indeed, in *statements* it is a common device for signalling, notably, continuation or, at the level of discourse, turn-keeping (Dutch: 't Hart 1998, Caspers 1998, 2000a, van Donzel 1999). Rather, for the intonational expression of the interrogativity function, speakers of Dutch appear to draw on the overall set of phonological devices of their language (be it that certain configurations are more likely to occur in questions than in statements).

Of considerably greater import were the effects of interrogativity on the *phonetics*. The results of the present study made it clear that the characteristic acoustic features of question intonation mainly arise in the phonetic implementation. Crucially, these phonetic features are taken to convey certain abstract meanings, that

is, what is being expressed by means of intonation is not uniquely signalled by the linguistic, phonological categories of a language. Rather, as proposed by Gussenhoven (1999b, 2002 to appear), the implementation of these categories offers the phonetic space for expressing additional *non*-linguistic meanings which derive from universal, innate codes such as the Frequency Code and the Effort Code (the latter associates larger pitch movements with greater significance or emphasis, Gussenhoven 2002 to appear). Such gradient universal meanings may 'enrich' the abstract meanings of phonological categories; their effects could be widely observed in our data. Thus, in the questions the low initial boundary tones (%L) had a systematically higher f0 than in the corresponding statements; the same held for the minima and for the postnuclear lows[1]. Likewise, between the statements and the questions the phonetic implementation of the H*(L) pitch accents showed substantial differences. Also, the three question types displayed systematic differences in the phonetic implementation of the final rise (H%). By itself, the presence of a final rise agreed with prior intuitions and claims that, in questions, H% is a major acoustic attribute. But over and above that, the final rises in DQ (whose question status has to be conveyed exclusively by prosodic means) were realised highest in the overall range, in WQ lowest. These findings suggested that phonetic implementation may play a major role in conveying certain abstract, universal meanings.

In the event of an affirmative reply to the first of the above research questions, two subquestions presented themselves:

- Do the observed intonational properties of interrogativity vary as a function of the *type* of question they occur in?
- Do the observed intonational properties of interrogativity vary as a function of the *speaker's sex*?

The first subquestion received an unambiguously affirmative answer. In chapter 1, three question types were selected for investigation. Of these, the wh-question (WQ) has two non-acoustic markers of interrogativity, a question word and inversion, the yes-no question (YQ) features only inversion, and the declarative question (DQ) has neither. We assumed that the relative importance of intonation in questions is greater when question word and/or inversion are absent. Accordingly, we formulated the Functional Hypothesis predicting an inverse relationship between the strength of acoustic features and the presence of non-acoustic markers of interrogativity. Since we also expected questions to feature higher pitch than statements (an expectation deriving from the Frequency Code), we predicted that pitch would be increasingly higher in the order WQ<YQ<DQ (with '<' representing 'more high/rising pitch'). As will presently be seen, this proved correct for almost all individual parameters (see below). Hence, it could be concluded that the high(er) pitch levels accompanying question intonation varied as a function of the type of question they occurred in.

[1] In the statements, the postnuclear low (PL) was the utterance-final low f0, in the questions, it constituted the low starting point of the final rise; cf. chapter 3, § 3.6.4.

The second subquestion could also be answered in the affirmative, although here some qualifications need to be made. So far, no hard and fast evidence has been found that, after suitable normalisation, register span in female speech is wider than in male speech, although one would expect this to be the case given the widespread assumption that female speech is more expressive and more varied. However, studies investigating sex-based differences in span always involved *declarative* speech. Our expectation was that *if* male and female speech differ in register span, this would show up more readily in questions than in statements. This idea was based on empirical evidence that (i) female communicative behaviour is primarily geared to interaction whereas that of men aims at assertion and carrying out tasks, and that (ii) women are less afraid to show the dependence implied by asking questions. In fact, it has been suggested that, for women, asking questions may be part of a strategy to obtain information, or to get things done. Therefore, it seemed reasonable to expect female speakers to exaggerate the acoustic features signalling interrogativity. Furthermore, considering that earlier investigations of sex-based differences in register span had mostly concentrated on *global* measures, we intended to take *local* measures into account as well. That is, we predicted that the excursions of nuclear accents and of final rises would be larger in the women's questions than in those of the men.

Overall, we found a clear tendency in the female speech towards larger local excursions[2], although significance was reached only with respect to the final rise and only at the .05 level[3]. In the domain of global measures, the difference in global register span was significant, again only at the .05 level. Interestingly, we also found that, in the female renderings, there was greater acoustic differentiation between the four different utterance types than in the male renderings (see § 7.2.3.2 below). We interpreted this as support for existing claims that female speech is the more expressive of the two.

Following this general outline of the results, the next sections deal more specifically with the hypotheses designed to answer the central research questions.

7.2.2 Hypotheses

Basing ourselves on the acoustic manifestations of interrogativity reported for other languages, we hypothesised that Dutch questions would feature:

- A final rise
- A high(er)-pitched onset
- A globally raised register level
- A less declining or even inclining global trend
- A raised nuclear accent peak.

These hypotheses were successively tested; the results are considered below.

---

[2] Use was made of the ERB scale, permitting direct comparison of pitch intervals in male and female speech (cf. chapter 3, § 3.4.3).

[3] Since the number of analyses of variance that had to be carried out was large, the significance level had been set at .01 rather than .05.

7.2.2.1 Final rises

We regarded the final rise as a grammaticalisation of the biological Frequency Code. Originally a gradient iconic attribute of, e.g., questions, it must at a certain point have evolved into a discrete, phonological category (Gussenhoven 1999b:291) which is either present or absent. As noted earlier, the hypothesis that the questions, unlike the statements, would predominantly feature this phonological category received ample support (questions 87% vs. statements 0%). However, the individual percentages for the three question types (WQ: 64%, YQ: 96%, and DQ: 100%) suggested that it is only in DQ that H% is truly indispensible. In WQ, the incidence of final rises was relatively high, given claims that, in languages such as English and German, the final rise is a marked option. However, it is not altogether clear whether the primary function of the final rise in this question type is to signal interrogativity, or whether it offers an opportunity to express additional pragmatic meanings, such as politeness and friendliness. Claims to this effect have been made for various languages. The finding that women realised a final rise in 78% of WQ, against 50% in the men, may be interpreted either way. The female speakers may have aimed to be maximally explicit about the question status of the utterance, even if this was already conveyed by means of the question word and inversion. This would fit in with a view that women are more intent on making explicit use of questions than are men. An alternative view is that women may have produced more final rises with a view to adding shades of attitudinal meaning. Since women have been found to be more listener-directed and more oriented towards communion than men, their final rises in WQ may have expressed "greater regardfulness, a more sympathetic attitude towards the interlocutor' (as claimed for British English by Schubiger 1959, cited in Bolinger 1989:32[4]). On the basis of our data, this issue cannot be settled; further experiments in this area are called for.

7.2.2.2 High onsets

For the hypothesis that onsets in questions would be higher than in statements there was both paradigmatic and syntagmatic support. Paradigmatically, the mean onsets of WQ and YQ were significantly higher than the mean onsets of the statements. As for DQ, means were, likewise, higher but here the difference was not significant. The syntagmatic approach involved the calculation of the amount of reset in utterance pairs with a statement in first position and a target utterance in second position; the latter could be ST, WQ, YQ, or DQ. In each pair, we compared both the onset value and the offset value of the initial statement with the onset value of the subsequent target utterance, on the idea that, between statements and questions, pitch would be reset to a higher level than between two consecutive statements. In the onset-onset comparisons, the results replicated those of the paradigmatic approach, that is, the onsets of WQ and YQ were significantly higher than the

[4] Note that is a clear connection between these attitudes (which reflect a certain degree of submissiveness), and higher pitch as predicted by the universal Frequency Code; see § 7. 2.3.4 below.

onsets of ST, whereas those of DQ were not. In the offset-onset comparisons, all three question types displayed significantly higher reset levels than the statements.

Rather unexpectedly, it was only in 11% that the systematically higher onsets in the questions corresponded with phonologically high or high-falling initial boundary tones (%H: 8%, %HL: 3%). Apparently, these phonological categories are not obligatory as long as the onsets of statements and questions are differentiated by the phonetics. This seems to accord well with Gussenhoven's proposal that phonological categories leave enough phonetic space for the additional expression of gradient universal meanings (cf. § 7.2.1).

7.2.2.3 Raised register level

Since, in numerous languages, interrogativity has been reported to cause f0 minima to be raised, we hypothesised that this would be the case in Dutch, too. Our data confirmed this. Not only did the questions display higher onsets than the statements, their f0 minima were, likewise, significantly higher. Indeed, we assumed that the high onset was, in fact, a corollary of an overall raised register level, rather than a parameter in its own right. Presumably, raised onset and raised minima in questions act as (early) indicators of the specific discoursal status of the utterance. This may be quite effective in spontaneous conversation given that, in this speaking style, replies have been found to overlap with the final rises of preceding questions (English: Geluijkens 1988). Such loss of the final rise (which, after all, is an important perceptual cue to questionhood) may well be compensated for by an overall raised register level. Nonetheless, assigning phonological status to raised register level seemed unwarranted. Unlike H%, which is categorically present or absent, a register shift is gradient and hard to identify as a discrete event. Rather, it is assumed that by raising (or lowering) the register level speakers express gradient, universal meanings (deriving, in this case, from the Frequency Code, see § 7.2.3.4 below). This, then, appears to make manipulation of register level a general phonetic device for lending greater prototypicality to utterances that express certain speaker attitudes or emotions, such as questions.

7.2.2.4 Global trend (declination vs. inclination)

For a comparison of global trends, register span and register level across the four utterance types, use was made of semi-automatically[5] fitted upper and lower regression lines, rather than of hand-fitted declination lines. According to the hypothesis, the global f0 trends (notably the lower regression slopes) would show less declination in questions than in statements; some question types, notably DQ, might even display inclination (cf. the Functional Hypothesis). However, mean slopes did not show a categorical clustering of statements versus questions. Instead, in WQ downslopes were steeper than in ST. This cast doubt on the possibility that global slope represents an independent speaker choice, with the most downward slope corresponding to statements, and less steep or even upward slope to questions (as reported for Danish). In chapter 4, it became apparent that the regression slopes

[5] Prior to the calculation of the regression lines, the final rises had to be identified and eliminated.

largely reflected the patterns of *local* pitch accents and postnuclear lows typical of each utterance type. For a start, the four utterance types differed phonologically as to the presence/absence of pitch accents on the subject (in the questions, these were frequently lacking), and the presence/position of high plateaus (which, in the questions, occurred relatively frequently).

More important, however, were the effects of differences in phonetic realisation. Thus, as noted before, in the questions mean f0 of %L was higher than in the statements. Similarly, in the questions subject accents were mostly reduced, whereas object accents were relatively large, causing pitch between subject and object accent to step up (see below). Moreover, the circumstance that the largest accent in WQ (on the wh-phrase) mostly occurred utterance-initially, whereas the largest accent in YQ/DQ (on the object) occurred utterance-finally, strongly affected the overall trends. Finally, global trends also co-varied with the postnuclear lows: in the questions featuring a final rise, this value was significantly higher than in questions lacking a final rise. In sum, we found no evidence that global trend constitutes an independent speaker choice. Rather, differences in slope appeared to arise from differences in local phonological and, notably, local phonetic properties of the utterances.

7.2.2.5 Raised nuclear accent peak

The hypothesis that questions would feature a raised nuclear accent peak (in our material: the accent on the object) was supported by the results. Whereas, in the statements, the accents on subject and object were roughly equal, pitch in YQ and DQ showed a strong tendency to step up between these two accents. Statistical analyses of the accents along eight parameters showed that this asymmetric pattern was, typically, due to a reduction of the excursion of the subject accent and a raised peak in the object accent; differences with ST were significant. As for WQ, comparison with the other utterance types on this count was not really possible: in most renderings, the subject-object sentence was pragmatically and prosodically subordinate to the wh-phrase.

7.2.2.6 Functional Hypothesis

Overall, our data met the expectation that the share of high/rising question pitch would be greatest in the question type lacking the lexical and syntactic properties of interrogativity (DQ), and smallest in the question type featuring both a question word and inversion (WQ). That is, results generally showed the order WQ<YQ<DQ (though not all pairwise differences in the post-hoc analyses of the Manova's proved significant[6]). However, DQ deviated from the hypothesised order by patterning with ST on two counts. This raised the question of whether this question type might be better suited for asking confirmation than information, a possibility hinted at in the literature. To test this hypothesis, a pen-and-paper experiment was carried out in chapter 5, the results of which are summarised in § 7.2.3.1 below.

[6] In all, 17 Manova's and 102 post hoc analyses were carried out.

7.2.3 Additional issues

As noted in the Introduction, our data sparked off a number of additional observations that had not been explicitly anticipated. Below, these are successively considered.

7.2.3.1 DQ: information question or confirmation question?

When selecting question types for the present study, we assumed that questions asked to elicit information are more prototypical than questions seeking confirmation. Following claims in the literature, we took it for granted that YQ and DQ are both information questions, with DQ merely being a subtype of YQ. That is, neither in YQ nor in DQ can the speaker predict whether the reply will be *yes* or *no*: basically, the information elicited by the question is new to him. On two acoustic variables, however, DQ proved somewhat closer to the statements than to YQ, suggesting that there might be some truth in hints in the literature that DQ is, in fact, biased towards confirmation. In a subsequent pen-and-paper experiment, this bias was operationalised as the degree of predictability (for the speaker) of the question's reply, such that '0%' stood for 'the reply is maximally unpredictable' and '100%' for 'the reply is maximally predictable'. Subjects rated written questions on this predictability scale by ticking percentages between 0 and 100.

As expected, replies to YQ were judged least predictable, i.e. *yes* and *no* were thought to be equally probable (mean predictability score: 2.6%). By contrast, the mean predictability score for DQ was 31.6%, reflecting a significant difference with YQ. At the same time, however, the presence of a particle and/or tag typically associated with confirmation raised the mean predictability scores to 84%. That is, the presence of *dus* or *hè?* had a stronger effect on predictability than question type. So, in the event DQ occupied an intermediate position on the predictability continuum: its subsequent reply was judged significantly more predictable for the speaker than a reply to YQ, but also considerably less predictable than replies to questions with *dus* and/or *hè?*.

The results of this experiment clearly established the degree in which the two question types are functionally different from each other. Also, they offered a possible explanation for the observed acoustic differences between YQ and DQ.

7.2.3.2 Upsweep

In all three question types, means of the utterances' onsets were phonetically higher than in the statements, in the order DQ<YQ<WQ. At the same time, the final rises of WQ were realised lowest in speakers' ranges, those of DQ highest; that is, here the order was reversed into WQ<YQ<DQ. Nonetheless, even in WQ with its relatively high utterance onset and its relatively low-range final rise, speakers generally made the utterance offset higher than the utterance onset (chapter 3, §

6.1). This suggested a strategy to give questions an upward slant, such that the final frequency is the highest of the entire utterance[7]. This phenomenon was termed UPSWEEP and it appeared fairly robust. Interestingly, in the female renderings of WQ, YQ and DQ, the amount of upsweep (i.e. the difference between utterance-final frequency and utterance-initial frequency) was substantially larger than in the male tokens. In the tokens *without* final rises (mostly statements), the difference between utterance-final f0 and utterance-initial f0 obviously yielded negative values ('downdrop'), presumably the joint result of declination and final lowering. The amount of downdrop was, similarly, greater in the female than in the male speakers. As a result, the female realisations of the four utterance types spanned a wider area of the overall pitch range. The observation that the women combined larger upsweeps with larger downdrops was taken to support claims that female speech is more expressive.

7.2.3.3 High plateaus

The questions featured relatively many high-pitched plateaus (cf. chapters 4 and 5), which could occur in several positions. Thus, plateaus were observed between the wh-word and the object accent (in WQ, subjects were accented only rarely), between subject accent and object accent, and between object accent and high final boundary tone. From a formal point of view, it seemed attractive to analyse all plateaus as products of the phonological adjustment rule DELETION, proposed by Grabe (1998; cf. Gussenhoven's Tone Linking Rule, 1984a), and extended here to include the final boundary tone H%. That is, before H% the second tone of an H*L or L*H pitch accent might be taken to be deleted. However, considering that the high(er) pitch levels in questions were assumed to derive from the Frequency Code (see § 7.2.3.4 below), an analysis claiming that speakers apply DELETION in order to raise overall pitch had to be rejected, since the rule works both ways. Thus, while deletion of L from H*L before H% leads to an increase of the overall proportion of high(er) pitch, deletion of H from L*H results in a longer stretch of *low* pitch. These opposite effects made it unlikely for the Frequency Code to have motivated application of an extended DELETION rule. Results of a perception experiment (Haan, Heijmans, Rietveld & Gussenhoven 2000) confirmed this. Listeners judging the pragmatic/attitudinal meanings of sets of (contextless) stimuli appeared to be chiefly guided by the high-low dichotomy. Thus, their scores were fully in line with FC-based predictions that, for some of the pragmatic scales, high mean pitch in the stimuli would be the most appropriate, whereas for the other scales the stimuli should, ideally, have low mean pitch. In other words, the structural effects of an extended DELETION rule on the pitch accents H*L and L*H did not correspond with a constant pragmatic/attitudinal effect. Instead, judgements appeared to be directly motivated by the biological Frequency Code. Relating these perception results to the high plateaus we assumed that, in the production of questions, the overall effect of the FC is that it prompts speakers to maximalise high pitch. One way of

[7] It may be added that the mean values of the final rises' maxima also exceeded the mean values of the (preceding) object peaks, in the male as well as the female speakers.

accomplishing this is by deleting L tones, a *phonological* operation resulting in high plateaus.

7.2.3.4 Operation of universal codes

If iconicity plays a more prominent role in intonation than in other areas of speech, our data suggested that this is especially true of question intonation. The main acoustic property of our question data, more high/rising pitch, was fully in line with predictions based on the universal Frequency Code (Ohala 1993, 1984, 1994). On the whole, this high(er) question pitch arose during the phonetic implementation. This fitted in well with Gussenhoven's claim (1999b, 2002 to appear) that the implementation of phonological categories offers scope for the gradient expression of universal meanings. With respect to the asymmetric accent patterns observed in the questions, we assumed the operation of another universal biological code, the Effort Code (Gussenhoven 1999b, 2002 to appear). This code, which associates larger pitch movements with greater discoursal significance, was taken to underlie focus structure. That is, the phonological category 'accent' may be variously implemented, such that the excursion of an accent heading a relatively significant constituent is larger than the excursion of an accent heading a constituent of relative insignificance.

Thus, interestingly, on the whole our question data appeared to invite a universalist analysis in which the observed form-meaning relations are not attributed exclusively to structural components, but also to paralinguistic and extralinguistic components. In our material, the acoustic differences between statements and questions were mainly gradient and phonetic, except for the final rise and the high plateaus which were regarded as categorical/phonological. Both these types of difference, the phonetic as well as the phonological, matched predictions made on the basis of Frequency Code and Effort Code. We proposed that the phonetic differences were *directly* motivated by these two biological codes and implemented as part of the realisation of the phonological categories. As for the phonological category H% ('final rise'), we analysed this as an *indirect* manifestation of the Frequency Code which, at a certain point in time, has discretised (cf. Gussenhoven 1999b). At the same time, we assumed that the Frequency Code *directly* causes questioners to choose the category H%, or to apply the phonological rule which deletes an L tone and produces a high plateau.

7.2.3.5 Focus in questions

The observation that the questions, unlike the statements, predominantly displayed asymmetric accent patterns suggested some relation with focus structure. In *statements*, a constituent is brought in focus because the speaker wants to direct the listener's special attention to it. While, acoustically, focus is realised by means of a pitch accent, the actual size of this accent generally depends on the constituent's communicative importance. Accordingly, manipulating the relative sizes of consecutive accents will bring about clear shifts in pragmatic meaning (as illustrated

by the focus materials of, e.g., Liberman & Pierrehumbert 1984, Rump & Collier 1996).

We assumed that the overall function of focus in *questions* is, likewise, to direct the listener's attention towards the focused constituent. However, given the major discoursal difference between statements and questions we conjectured that, between these two, the *motives* underlying focus might differ. Hence, in chapter 6 we formulated the following additional research questions:

- Considering that focus always aims to highlight a constituent, do statements and questions differ as to the purpose that is served by this highlighting?
- If such a difference exists, does it account for the observed variation in the accent patterns of statements and questions?

Crucially, our approach to focus was based on the universal dichotomy which divides utterances into the (pragmatic) constituents *topic* and *comment*. In statements, the comment commonly adds new information to the topic; hence, it is always [+focus]. In questions, the comment is, likewise, [+focus] but this is not because it adds information to the topic. Rather, we proposed, in questions the comment reflects the speaker's presupposition that is the object of questioning. Thus, in YQ and DQ, a speaker utters a comment that is currently entertained by him and the correctness of which he wants to check. He does this by (implicitly) contrasting the uttered comment with a set of other comments that might also pertain to the topic in hand. In WQ, the comment is realised by the question word. This represents the speaker's presupposition that some extension can be made with respect to a topic sentence that is already taken for granted. While the word-class of the question word defines the range of this extension (e.g., *Who?* points to a set of persons, *Why?* to a set of reasons), it also indicates that the speaker wants to be provided with the *correct* referent from the presupposed set.

In sum, in statements and questions speakers appear to have different motives for highlighting the comment. This amounts to saying that different focus *types* are involved, the acoustic properties of which may be different, too. Given that comments in questions crucially involve the notion 'correct item from a set', we proposed that the appropriate focus type is FOCUS OF CONTRAST (cf. Chafe 1976). As many studies have shown, contrastively focused constituents feature relatively large pitch excursions with steep falls, whereas neighbouring accents tend to be strongly reduced. While the accent patterns of the questions typically displayed these phonetic features, those of the statements did not. This led to the account that the various accent patterns in the questions and statements rested, to a large extent, on differences in focus type. Thus, the right-asymmetric pattern in YQ and DQ was jointly produced by the reduced accent on the topic constituent (in our material: the subject) and the relatively large accent on the comment constituent (in our material: the object). In WQ, the left-asymmetric pattern was the combined result of the utterance-initial position of the comment (the wh-phrase), its relatively large and steeply falling accent, and the reduction (or absence) of the accents on subject and object.

In conclusion, either of the research questions received an affirmative answer. A consequence of this is that our earlier definition of what is to be understood by a question has to be slightly modified. Thus, rather than being *an utterance intended to function as a request for information* (cf. chapter 1), a question should be redefined as *an utterance intended to check the comment's correctness.*

7.3 Discussion: Theoretical and practical implications

In Dutch, interrogativity can be conveyed by lexical means (wh-word), syntactic means (inversion), gestural means (eyebrows, facial expression), as well as by various intonational means that may be phonological as well as phonetic. Obviously, this is true of most other languages and it is probably not without reason that, crosslinguistically, the encoding of interrogativity is so robust. Unlike statements, questions make a direct appeal to a listener for a response, and redundancy ensures that the listener identifies this specific discoursal status, even in difficult circumstances. The systematic investigations reported in this study have yielded a great deal of information about the *intonational* properties of Dutch questions. An important overall conclusion is that there is no such thing as 'question intonation', at least in the sense that the intonation of questions should require a specific overall melody or phonological structure. This is in keeping with claims to this effect by Gussenhoven (1984a:203[8]) and 't Hart (1998:103[9]), one of the authors of the IPO grammar of Dutch intonation (cf. chapter 2). These claims, however, did not include statements about possible *phonetic* properties of question intonation. What the results of the present study have made clear is that question intonation in Dutch is largely a matter of phonetics, and it seems likely that these phonetic features are instrumental in making a listener aware that he is expected to produce a reply. In our production material, the intonation of questions was made distinct from the intonation of statements by:

- Overall high(er) pitch which manifested itself chiefly in the phonetics, and
- A different focus type (FOCUS OF CONTRAST) which, likewise, manifested itself in the phonetics and produced asymmetric accent patterns.

As pointed out earlier, our data supported recent proposals by Gussenhoven (1999b, 2002 to appear) that the implementation of phonological categories offers the phonetic scope for expressing universal meanings deriving from innate, biological codes. As will be clear, such an analysis lends support to the view that there are direct, non-linguistic channels between prosody and interpretation bypassing

[8] "Observe that in the intonational system there is no marker for a category 'interrogative' comparable to subject-verb inversion in syntax". It may be noted that, in Gussenhoven's description, the appropriate intonation for "speech acts that can be described as 'requests for information'" arises from combinations of certain pitch accents (which have their own discoursal meanings) with a binary 'serving' parameter. The latter causes pitch accents to be 'speaker-serving' or 'hearer-serving', that is, they are employed for the benefit of the speaker or the hearer, respectively.

[9] "There is no specific intonation pattern characteristic of questions".

phonological structure. In the statements as well as in the questions, %L (low initial boundary tone) and H*L (falling pitch accent) were by far the commonest phonological categories; what chiefly varied as a function of discoursal category was their phonetic implementation. Also, though most questions shared the phonological category H%, the latter's phonetic implementation crucially varied as a function of question type[10]. Likewise, the intricacies of information structure (as reflected by focus structure) would also seem to be mediated by a non-linguistic channel. Thus, the phonetic variation in realisations of the linguistic category H*L, which signalled the relative significance of the corresponding constituents, was taken to be motivated by the non-linguistic Effort Code.

The results thus favoured a theoretical approach of intonation which, typically, incorporates universal, non-linguistic aspects. The overall dependence of question intonation on gradient, universal meanings would seem to make questioning an attitude rather than a linguistic category. Still, for the acoustic expression of this attitude languages may vary in the use they make of universal codes. Also, opinions differ with respect to the degree in which universal codes may affect the phonology of a given language. Lindsey (1985) claims that phonological changes will always be held in check by iconicity. That is, such changes will never violate the basic correlations between pitch and abstract iconic meaning. Gussenhoven (2002 to appear), however, shows convincingly that it is possible for phonological change to turn against the iconicity motivated by, e.g., the Frequency Code. For instance, in some languages specific phonological developments may have yielded falling forms in questions, rising forms in statements. Crucially however, universal codes may yet counteract such violations of 'naturalness' by asserting themselves in the phonetic implementation. To this, it may be added that the *choice* of (arbitrary) phonological category may, also, be iconically motivated. Thus, speakers expressing meanings or attitudes which the universal Frequency Code associates with high pitch may, preferably, select (arbitrary) H%, or decide to delete the L tone of (arbitrary) H*L.

In a more practical sphere, the present study has shown that regression lines may constitute a useful methodological tool for comparing overall trend, register level and register span across large numbers of utterances. When the data are phonologically homogeneous, regression lines may bring out systematic differences in phonetic implementation. Evidently, our results also have consequences for speech technology. In automatically generated speech, it is crucial for a question to be unequivocally *perceived* as a question. Clearly, an ensemble of acoustic cues to interrogativity occurring throughout the utterance is more likely to ensure a correct question interpretation than the mere addition of a H% at the utterance end. For this ensemble, our findings offer useful ingredients. Machine-generated questions should be given higher onsets and higher minima, and the highest f0 value of the final rise

[10] Gussenhoven & Chen (2000) showed that such variation in phonetic implementation is interpreted along the same (universal) lines by speakers of Chinese, Hungarian and Dutch, although these languages express interrogativity in different manners. Listening to stimuli from an allegedly little-known language, the subjects consistently selected the stimuli with the higher peaks and the higher final frequency as the ones that signalled questions, regardless of their own native language.

should exceed the f0 of the nuclear accent and of the utterance onset (UPSWEEP). When the question is multi-accented, the phonetic implementation of the accents should follow both from the focus distribution across topic and comment, and from focus type. Thus, in questions the accent on the topic should commonly be reduced or altogether absent, whereas the accent on the comment has to be made relatively salient (FOCUS OF CONTRAST). Implementation of these phonetic features in machine questions will no doubt improve the latter's effectiveness in man-machine dialogues. By the same token, our results may be useful for automatic speech recognition. In dialogues between man and machine, it is crucial for the latter to be able to distinguish questions from statements. This means that automatic systems may be made sensitive to the acoustic properties of questions identified by the present investigation.

7.4 Suggestions for further research

Drawing up a systematic account of the properties of the intonation of Dutch questions is a useful undertaking, and the present study should be seen as a first step towards this goal. Starting more or less from scratch, we concentrated on what seemed core aspects of Dutch question intonation, making use of read speech only. The huge advantage of this speech style is that it can be maximally controlled, enabling the researcher to pinpoint subtle differences and set up prototypical contours. An obvious next step is to test the robustness of such prototypes in spontaneous speech; this should be one of the prime objects of a future study. In point of fact, as a part of the present study we investigated 154 spontaneous utterances taken from doctor-patient consultations, a type of dialogue conducive to asking questions (Van Heuven, Haan & Pacilly 1998). Ten taped conversations between general practitioners (5 male, 5 female) and their patients (5 male, 5 female) yielded transcripts totalling two hours of interaction. On the basis of the written transcripts, questions were classified as WQ, YQ, DQ, Elliptic Questions, Echo Questions, and Tag Questions. However, due to the poor sound quality and to speech overlap, only 140 of the 234 questions could be acoustically analysed and compared with statements; of the latter utterance type, only 14 proved fit for analysis. Although, broadly, the results were in line with those obtained in the corpus of read speech, it was decided not to pursue the investigation given the lexical, durational and typological heterogeneity of the materials. Evidently, one of the constraints to be put on future experiments in the area of spontaneous questions is that the elicited data are more controlled.

Considering, further, that the present study was restricted to production data, perception experiments will be needed to verify the conclusions. A number of separate perception experiments have already been carried out as part of our investigations of Dutch question intonation; results were reported in Van Heuven, Haan, Janse & Van der Torre (1997), Van Heuven, Haan & Pacilly (1997), Van Heuven & Haan (2000a), Van Heuven & Haan (2000b) and Van Heuven & Haan (2002 to appear). In these experiments, listeners were presented with increasingly

larger speech fragments of statements and questions (gating experiments), their task being to indicate which utterance type they heard[11]. Source utterances for these gating tests were taken from the production corpus described in the present study. They were either used in their original form or they were resynthesised in order to manipulate acoustic variables (declination slope, peak heights of subject and object, presence/absence of final rise[12]). So far, the results made it clear that the acoustic characteristics emerging from our production data are relevant for perception, too. To give an example, an appreciable number of the stimuli lacking a final rise yet elicited 'question' responses. Apparently, when classifying these utterances as questions listeners were guided by other acoustic cues, such as declination rate[13] or excursion size of the nuclear accent (cf. Van Heuven & Haan 2000b, 2002 to appear). This confirmed the conclusion of the present production study that a final rise is not truly indispensable in questions.

Another area requiring further experimental research is that of the acoustic and pragmatic implications of the absence vs. presence of a final rise. Our data indicated that utterances featuring a final rise showed less global downtrend than utterances without H% (cf. chapter 3, § 6.4). The acoustic relationship between overall slope and a projected H% certainly seems worth further investigation. As for the pragmatic consequences of the presence vs. absence of a final rise, more tests should be carried out to uncover the range of attitudes that may be expressed in this way. In order to preclude possible interference from the Frequency Code[14], care should be taken to systematically vary mean pitch level, such that low mean pitch is combined with a final rise, and vice versa. Also, it should be investigated to what extent listeners' judgements correspond with predictions based on the Frequency Code.

Finally, as noted earlier, our approach to focus in questions appears fairly novel. So far, the knotty issue of focus in questions seems to have been largely evaded. In chapter 6 of this study, we made an attempt to come to grips with it by combining some isolated hints in the literature, i.e., an observation that wh-phrases constitute comments (Bolinger 1990), and a few brief speculations about a possible relation between questions and contrastiveness (Van Es 1932, Kiefer 1980, Eady & Cooper 1986). While it is important to stress the exploratory character of our present proposals on this subject, it seems equally important to argue for more production and perception experiments in this area, the results of which may shed more light on the feasibility of the approach outlined here. In particular, given that all our utterances featured *two* accents whose patterns gave rise to the present account,

[11] In some experiments, subjects had to settle on one type out of ST/WQ/YQ/DQ, in others they were presented with ST vs. DQ only.
[12] Register level was not manipulated. In future experiments, this parameter should be included, too.
[13] Obviously, the finding that declination rate serves as a perceptual cue does not imply that, conversely, speakers are intent on differentiating between overall slopes. As concluded in chapter 3 and 4, the different slopes in our material mainly resulted from *local* differences in position and f0 of onsets, accents and postnuclear lows.
[14] Cf. Haan et al 2000, which showed that judgements were inspired by mean pitch level rather than by morphological structure.

References

Arvaniti, A., D. R. Ladd and I. Mennen (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26, 3-25.

Athanasiadou, A. (1991). The discourse function of questions. *Pragmatics* 1, 107-122.

Bartels, C. (1999). *The intonation of English statements and questions. A Compositional Interpretation.* New York & London, Garland Publishing Inc.

Bartels, C. and J. Kingston (1994). Salient Pitch Cues in the Perception of Contrastive Focus. P. Bosch and J. van der Sandt (eds), *Focus and Natural Language Processing*, Vol. 1: *Intonation and Syntax.* IBM Working Papers of the Institute for Logic and Linguistics, 11-28.

Basow, S. A. (1992). *Gender: stereotypes and roles.* Pacific Grove CA, Brooks/Cole.

Beckman, M. E. (1995a). Problems of Intonation. *Proceedings of the XIIIth International Congress of Phonetic Sciences* Vol. 1, 450-457. Stockholm, KTH and Stockholm University.

Beckman, M. E. (1995b). Local shapes and global trends. *Proceedings of the XIIIth International Congress of Phonetic Sciences* Vol. 2, 100-107. Stockholm, KTH and Stockholm University.

Beckman, M. E. and J. B. Pierrehumbert (1986). Intonational structure in Japanese and English. *Phonology Yearbook* 3, 255-309.

Belnap, N. and T. B. Steel (1976). *The logic of questions and answers*. New Haven, Yale University Press.

Benkirane, T. (1998). Intonation in Western Arabic (Morocco). Hirst, D. and A. Di Cristo (eds), *Intonation Systems. A Survey of Twenty Languages*, 345-359.Cambridge University Press.

Berryman-Fink, C. and J. R. Wilcox (1983). A Multivariate Investigation of Perceptual Attributions Concerning Gender Appropriateness in Language. *Sex Roles*, Vol. 9, no.6, 663-681.

Beun, R. J. (1990). The recognition of Dutch declarative questions. *Journal of Pragmatics* 14, 39-56.

Biemans, M. (1998). The Effect of Biological Gender (Sex) and Social Gender (Gender Identity) on Three Pitch Measures. In: R.van Bezooijen en R. Kager (eds), *Linguistics in The Netherlands 1998*, 41-52, Amsterdam, John Benjamins.

Boersma, P. and D. J. M. Weenink (2000). Praat: Doing phonetics by computer. http://www.fon.hum.uva.nl/praat/

Bolinger, D. (1972a). Relative Height. D. J. Bolinger (ed.), *Intonation*. Hammondsworth, Penguin, 137-155.

Bolinger, D., (1972b). Accent is predictable (if you're a mind-reader). *Language 48*, 633-644.

Bolinger, D. (1978a). Asking more than one thing at a time. H. Hiz (ed.), *Questions*, 107-150. Dordrecht/Boston, Reidel.

Bolinger, D. (1978b). Intonation across languages. J. Greenberg (ed.), *Universals of Human Language* Vol. II: Phonology, 471-524. Palo Alto (Calif.), Stanford University Press.

Bolinger, D. (1982). Nondeclaratives from an Intonational Standpoint. R. Schneider, K Tuite and R. Chametzky (eds), *Papers from the Parasession on Non-declaratives* 18, 1-22. Chicago Linguistic Society.

Bolinger, D. (1986). *Intonation and Its Parts.* London, Edward Arnold.

Bolinger, D. (1989). *Intonation and Its Uses*. Palo Alto (CA), Stanford University Press.

Bolinger, D. (1990). Accent in prototypical wh-questions. S. L. Tsobatzidis (ed.), *Meanings and prototypes: Studies in linguistic categorization*, Chapter 14. London, Routledge.

Brend, R. (1975). Male-female intonation patterns in American-English. B. Thorne and N. Henley (eds), *Language and Sex:Difference and Dominance*. Rowley (MA), 84-87. Newbury House.

Brown, G., K. Currie and J. Kenworthy (1980). *Questions of Intonation.* London, Croom Helm.

Bruce, G. (1977). *Swedish word accents in sentence perspective.* Lund, Gleerup.

Bruce, G. (1982a). Textual Aspects of Prosody in Swedish. *Phonetica* 39, 274-287.

Bruce, G. (1982b). Developing the Swedish intonation model. *Working Papers Department of Linguistics, University of Lund*, nr. 22, 51-116.

Caspers, J. (1996). Testing the semantics of single-accent intonation patterns. A. Arregui and C. Cremers (eds), *Meaning on the HIL.* HIL Occasional Papers 1, 14-32. Leiden, Holland Institute of Generative Linguistics.

Caspers, J. (1997). Testing the meaning of four Dutch pitch accent types. G. Kokkinakis, N. Fakotakis and E. Dermatas (eds), *Proceedings of the 5th European Conference on Speech Communication and Technology,* Vol. 2, 863-866. University of Patras.

Caspers, J. (1998). Who's Next? The Melodic Marking of Question vs. Continuation in Dutch. *Language and Speech* 41, 375-398.

Caspers, J. (1999). The 'early' and the 'late' accent-lending fall in Dutch. R. van Bezooijen and R. Kager, *Linguistics in the Netherlands 1999*, 27-39. Amsterdam, J. Benjamins.

Caspers, J. (2000a). Looking for melodic turn-holding configurations in Dutch. H. de Hoop and T. van der Wouden (eds), *Linguistics in the Netherlands 2000,* 45-55. Amsterdam, J. Benjamins.

Caspers, J. (2000b). Pitch accents, boundary tones and turn-taking in Dutch map task dialogues. *Proceedings International Conference on Spoken Language Processing 2000*, Vol I, 565-568.

Caspers, J. Van Heuven, V. and N. Van Zwol (1998). An experimental investigation of meaning differences between the "pointed hat" contour and the accent-lending fall in Dutch. R. Van Bezooijen and R. Kager (eds), *Linguistics in the Netherlands 1998,* 65-79. Amsterdam, Benjamins.

Cauldwell, R. (1999). Judgements of attitudinal meanings in isolation and in context. *Phonetics Teaching & Learning Conference 1999.* London, University College.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects and topics. C. Li (ed.), *Subject and topic*, 25-56. New York, Academic Press.

Chen, X.S. (1990). *The Prosody of Mandarin Chinese.* Berkeley, University of California.

Chisholm Jr., W. S. (1982) (ed.). *Interrogativity: a colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam, J. Benjamins.

Clements, G. N. (1990). The status of register in intonation theory: comments on the papers by Ladd and by Inkelas and Leben. J. Kingston & M. Beckman (eds), *Papers in Laboratory Phonology I*, 58-71. Cambridge University Press.

Cohen, A., R. Collier and J. 't Hart (1982). Declination: Construct or Intrinsic Feature of Speech Pitch? *Phonetica* 39*,* 254-273.

Collier, R. (1987). F0 Declination: The Control of Its Setting, Resetting, and Slope. T. Baer, C. Sasaki and K. Harris (eds), *Laryngeal Function in Phonation and Respiration*, 403-421. Boston, College-Hill Press.

Collier, R. (1989). On the phonology of Dutch intonation. F. J. Heyvaert and F. Steurs (eds), *Worlds behind words,* 245-258. Leuven University Press.

Collier, R. (1993). On the communicative function of prosody: some experiments. *IPO Annual Progress Report* 28, 67-75. Eindhoven.

Couper-Kuhlen, E. (1984). A New Look at Contrastive Intonation. R. J. Watts and U. Weidmann (eds), *Modes of Interpretation: Essays presented to Ernst Leisi.* Tübingen, Gunter Narr, 137-158.

Couper-Kuhlen, E. and M. Selting (1996). Towards an interactional perspective on prosody and a prosodic perspective on interaction. M. Selting and E. Couper-Kühlen (eds), *Prosody in conversation: interactional studies.* Cambridge University Press, 11-56.

Cruttenden, A. (1981). Falls and rises: meanings and universals. *Journal of Linguistics* 17*,* 77-91.

Cruttenden, A. (1992). Review of 't Hart, Collier and Cohen 1990, *Journal of Linguistics* 28*,* 522-526.

Cruttenden, A. (1994), *Intonation*. Cambridge University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English.* Cambridge University Press.

Cutler, A. (1977). The context-dependence of "intonational meanings". *Papers from the 13th Regional Meeting,* Chicago Linguistic Society, 104-137.

Cutler, A., D. Dahan, W. van Donselaar (1997). Prosody in the Comprehension of Spoken Language: A Literature Review. *Language and Speech* 40 (2), 141-201.

Daan, J. (1938). Dialect and Pitch Pattern of the Sentence. *Proceedings 3rd International Congress of Phonetic Sciences* Ghent, 473-481.

Dascalu-Jinga, L. (1998). Intonation in Romanian. Hirst, D. and A. Di Cristo (eds), *Intonation Systems. A Survey of Twenty Languages*, 239-260. Cambridge University Press.

De Groot, A.W. (1943). De Nederlandse zinsintonatie in het licht der structurele taalkunde. *De Nieuwe Taalgids 37,* 30-41.

De Groot, A.W. (1949/1965), *Structurele Syntaxis*. Den Haag, Servire.

Den Hertog, C.H. and H. Hulshof (1972). *Nederlandse Spraakkunst*. Amsterdam, W. Versluys.

D'Imperio, M. (1997). Breadth of focus, modality and prominence perception in Neapolitan Italian. K. Ainsworth-Darnell and M. D'Imperio (eds), *OSU Working papers in Linguistics 50*, 19-39.

D'Imperio, M. and D. House (1997). Perception of questions and statements in Neapolitan Italian. *Proceedings ESCA Eurospeech97 Rhodes* Vol. 1. University of Patras, 251-254.

Di Cristo, A. (1998). Intonation in French. D. Hirst and A. Di Cristo (eds), *Intonation Systems, a Survey of Twenty Languages*, 195-218. Cambridge University Press.

Dik, S. *et al.* (1980). On the typology of Focus Phenomena. *GLOT* (3-4), Themanummer Perspectives on Functional Grammar, 41-74.

Droste, G. (1972). Vragen met "he". *Leuvense Bijdragen* 61, 121-133.

Eady, S. J. and W. E. Cooper (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80 (2), 402-415.

Eady, S. J., W. E. Cooper, G. V. Klouda, P. R. Müller and D. W. Lotts (1986). Acoustical characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments. *Language and Speech* 29 (3), 233-251.

Es, G.A. van (1932). Syntaxis en dialectstudie IV: Intonatie en syntaxis 2. *Onze Taaltuin* 4, 122-128, 169-175.

Fujisaki, H. (1983). Dynamic characteristics of voice fundamental frequency in speech and singing. P. MacNeilage (ed.) *The production of speech.* Heidelberg, Springer, 39-55.

Gårding, E. (1982). Prosodic expressions and pragmatic categories. *Working Papers Dept. of Linguistics*, Lund, 117-135.

Gårding, E. (1983). A generative model of intonation. A. Cutler and D. R. Ladd (eds) *Prosody:Models and Measurements.* Berlin, Springer, 11-25.

Geluykens, R. (1987). Intonation and Speech Act Type. *Journal of Pragmatics* 11, 483-494.

Geluykens, R. (1988). On the myth of rising intonation in polar questions. *Journal of Pragmatics* 12, 467-485.

Geluykens, R. (1989). R(a)ising questions: question intonation revisited. A reply to Batliner and Oppenrieder. *Journal of Pragmatics* 13, 567-575.

Givón, T. (1984). The speech-act continuum. W. S. Chisholm Jr. (ed.), *Interrogativity: a colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam, J. Benjamins, 245-254.

Goldsmith, J. (1976). *Autosegmental Phonology*. PhD thesis, MIT. Distributed by IULC and published 1979 by Garland Press, New York.

Gooskens, C. and V. J. van Heuven (1995). Declination in Dutch and Danish: global versus local pitch movements in the perceptual chracterisation of sentence types. *Proceedings of The XIIIth International Congress of Phonetic Sciences*, Stockholm, Vol. 2, 374-377.

Gordon, M. (1999). The intonational structure of Chickasaw. Proceedings 14th International Conference of Phonetic Sciences, San Francisco, 1993-1996.

Gósy, M. and J. Terken (1994). Question marking in Hungarian: timing and height of pitch peaks. *Journal of Phonetics* 22, 269-281.

Grabe, E. (1998). *Comparative Intonational Phonology: English and German.* Doctoral dissertation Max Planck Institute für Psycholinguistik and University of Nijmegen.

Grabe, E., C. Gussenhoven, J. Haan, E. Marsi and B. Post (1997). Preaccentual Pitch and Speaker Attitude in Dutch. *Language and Speech* 41, 63-85.

Greenberg, J., C.A. Ferguson and A. Moravcsik (1978). *Language universals.* Stanford (CA), Stanford University Press.

Grice, M. (1995). *The intonation of interrogation in Palermo Italian.* Tübingen, Niemeyer.

Grice, M. R. Benzmüller, M. Savino and B. Andreeva (1995). The intonation of queries and checks across languages: data from map task dialogues. *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, Vol. 3, 648-651.

Grice, M. and M. Savino (1997). Can pitch accent type convey information status in yes-no questions?. *Proceedings Workshop Concept-to-Speech Generation Systems, 29-38. ACL/EACL*, Madrid.

Groenendijk, J. and M. Stokhof (1996). Questions. J. van Benthem and A. ter Meulen (eds), *The Handbook of Logic and Language*, 1057-1124. Amsterdam, Elsevier.

Grønnum, N. (1992). *The groundworks of Danish intonation.* Odense, Museum Tusculanum Press.

Grønnum, N. (1995). Superposition and subordination in intonation, a non-linear approach. *Proceedings 13th International Congress of Phonetic Sciences,* Vol. 2, 124-131

Grønnum, N. (1998). A critical remark on D. R. Ladd's *Intonational Phonology*. *Journal of Phonetics* 26, 109-112.

Guittart, L. J. (1925). *De Intonatie van het Nederlands, met inbegrip van een Vergelijking met de Engelse Intonatie.* Utrecht, Kemink & Zoon.

Gundel, J. K. (1977). *Role of Topic and Comment in Linguistic Theory.* PhD dissertation, Indiana University Linguistics Club, Bloomington, Indiana.

Gundel, J. K. (1988). Universals of topic-comment structure. M. Hammond (ed.) *Studies in Syntactic Typology*, 209-239. Amsterdam, Benjamins.

Gundel, J. K. (1999). On Different Kinds of Focus. P. Bosch and R. van der Sandt (eds), *Focus. Linguistic, Cognitive, and Computational Perspectives.* Cambridge University Press, 293-305.

Gussenhoven, C. (1984a). *On the grammar and semantics of sentence accents.* Dordrecht, Foris.

Gussenhoven, C. (1984b). Deze stijging is dus een daling: een fonologische interpretatie van een fonetische beschrijving van de intonatie van het Nederlands. *Handelingen van het 38e Nederlandse Filologen Congres*, 303-321

Gussenhoven, C. (1985). Two views of accent: a reply. *Journal of Linguistics* 21, 125-138.

Gussenhoven, C. (1988a). Adequacy in Intonation Analysis: The Case of Dutch. H. van der Hulst and N. Smith (eds), *Autosegmental studies on pitch accent*, 95-121. Dordrecht, Foris.

Gussenhoven, C. (1988b). Intonational phrasing and the prosodic hierarchy. W. U. Dressler *et al.* (eds) *Phonologica 1988*, 89-99. Cambridge University Press.

Gussenhoven, C. (1991). Tone Segments in the Intonation of Dutch. T. F Shannon and J. P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1989*, 139-155. Lanham, University Press of America.

Gussenhoven, C. (1992a). Sentence accents and argument structure.. I. M. Roca (ed.), *Thematic Structure Its Role in Grammar*, 79-106. Berlin, Foris.

Gussenhoven, C. (1992b). Review of A perceptual study of intonation by J. 't Hart, R. Collier and A. Cohen. *Language* 68 (3), 610-614.

Gussenhoven, C. (1999a). On the Limits of Focus Projection in English. P. Bosch and R. van der Sandt, *Focus. Linguistic, Cognitive, and Computational Perspectives*, 43-55. Cambridge University Press.

Gussenhoven, C. (1999b). Discreteness and gradience in intonational contrasts. *Language and Speech* 42 (2-3), 283-305.

Gussenhoven, C. (2000a). The lexical tone contrast of Roermond Dutch in Optimality Theory. M. Horne (ed.), *Prosody: Theory and Experiment*, 129-167.Dordrecht, Kluwer.

Gussenhoven, C. (2000b). The Boundary Tones are coming: On the Nonperipheral Realisation of Boundary Tones. M. B. Broe and J. B. Pierrehumbert (eds) *Papers in Laboratory Phonology V*, 132-151. Cambridge University Press.

Gussenhoven, C. (2002a, to appear). Why question intonations rise and why they sometimes fall. M. Hoey and N. Scott (eds), *Questions: Multiple Perspectives on a Common Phenomenon* (to appear). University of Liverpool Press (Language and Discourse Series).

Gussenhoven, C. (2002b, to appear). Transcription of Dutch Intonation (ToDI). Sun-Ah Jun (ed.), *Prosodic Typology and Transcription: A Unified Approach* (to appear). Oxford University Press.

Gussenhoven, C. and A. Chen (2000). Universal and language-specific effects in the perception of question intonation. *Proceedings International Conference on Spoken Language Processing,*II, 91-94.

Gussenhoven, C., B. H. Repp, A. Rietveld, H. H. Rump and J. Terken (1997). The perceptual prominence of fundamental frequency peaks. *Journal of the Acoustical Society of America 102 (5) Pt. 1,* 1-14.

Gussenhoven, C. and T. Rietveld (1991). An experimental evaluation of two nuclear-tone taxonomies. *Linguistics* 29, 423-449.

Gussenhoven, C. and T. Rietveld (1992). A Target-Interpolation Model for the Intonation of Dutch. *Proceedings International Conference of Spoken Language Processing* Vol. 2, 1235-1238.

Gussenhoven, C. and T. Rietveld (1997). Empirical evidence for the contrast between H* and L* in Dutch rising contours. A. Botinis, G. Kouroupetroglou and G. Caryannis (eds), *Intonation: Theory, models and applications.* Proceedings ESCA Workshop, 169-172. University of Athens.

Gussenhoven, C. and T. Rietveld (1998). On the speaker-dependence of the perceived prominence of F_0 peaks. *Journal of Phonetics* 26, 371-380.

Gussenhoven, C. and T. Rietveld (2000). The Behavior of H* and L* Under Variations in Pitch Range in Dutch Rising Contours. *Language and Speech* 43 (2), 183-203.

Gussenhoven, C. and P. van der Vliet (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics* 35, 99-135.

Guy, G., B. Horvath, J. Vonwiler, E. Daisley, and I. Rogers (1986). An Intonational Change in Progress in Australian English. *Language in Society* 15, 23-52.

Haan, J., L. Heijmans, A. Rietveld and C. Gussenhoven (2000). Phonological undershoot before H% in Dutch intonation contours?. Poster presented at the 7th International Conference on Laboratory Phonology, Nijmegen.

Haan, J. and V. J. van Heuven (1999a). Male vs. female pitch range in Dutch questions. *Proceedings of the International Congress of Phonetic Sciences*, San Francisco, 1581-1584.

Haan, J. and V. J. van Heuven (1999b). Mannen, vrouwen en vragen: verschillen in toonbereik?. E. Huls and B. Weltens (eds), *Artikelen van de Derde Sociolinguistische Conferentie*, 149-160. Delft, Eburon.

Haan, J., V. J. van Heuven, J. Pacilly and R. van Bezooijen (1997). An anatomy of Dutch question intonation. J. Coerts and H. de Hoop (eds), *Linguistics in the Netherlands 1997*, 97-108. Amsterdam, Benjamins.

Haan, J., V. J. van Heuven, J. Pacilly and R. van Bezooijen (1997). Intonational characteristics of declarativity and interrogativity in Dutch: A comparison. A. Botinis, G. Kouroupetroglou and G. Caryannis (eds), *Intonation: Theory, models and applications.* Proceedings ESCA Workshop, 173-176. University of Athens.

Hadding-Koch, K. (1961). *Acoustico-phonetic studies in the intonation of southern Swedish.* Lund, Gleerup.

Hadding-Koch, K. and M. Studdert-Kennedy (1964). An Experimental Study of Some Intonation Contours. *Phonetica* 11, 175-185.

Haeseryn, W. , K. Romijn, G. Geerts, J. de Rooij en M.C. van den Toorn (1997). *Algemene Nederlandse Spraakkunst*. Groningen, M. Nijhoff.

Hajicová, E. (1983). On some aspects of presuppositions in questions. F. Kiefer (ed.), *Questions and Answers*. Dordrecht, Reidel, 85-96.

Halliday, M. A. K. (1967a). *Intonation and Grammar in British English.* The Hague, Mouton.

Halliday, M. A. K. (1967b). Notes on transitivity and theme in English (Part 2). *Journal of Linguistics* 3, 199-274.

Hayes, B. (1993). 'Gesture' in prosody: comments on the paper by Ladd. P. Keating (ed.), *Papers in Laboratory Phonology*, Vol. III: *Phonological structure and phonetic form*, 64-75. Cambridge University Press.

Henton, C. (1989). Fact and fiction in the description of female and male pitch. *Language and Communication* Vol. 9 No. 4 , 299-311.

Hermann, E. (1942). *Probleme der Frage*, 2 Vols. Göttingen, Vandenhoeck & Ruprecht.

Hermes, D. (1988). Measurement of Pitch by Subharmonic Summation. *Journal of the Acoustical Society of America* 83, 257-264.

Hermes, D. and J.C.van Gestel (1991). The Frequency Scale of Speech Intonation. *Journal of the Acoustical Society of America* 90, 97-102.

Hetland, J. and V. Molnár (1995). Zur Funktion der Prosodie in Fragen. *Sprache und Pragmatik,* Sonderheft (Festvorträge anläßlich des 60. Geburtstag von Inger Rosengren, Lund), 3-46.

Higginbotham, J. (1995). The Semantics of Questions. S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, 361-383. London, Blackwell.

Hirschberg, J. (1989). Distinguishing Questions by Contour in Speech Recognition Tasks. *Proceedings Speech and Natural Language Workshop DARPA, 22-35*. San Mateo (Calif.), Kaufman.

Hirst, D. and A. Di Cristo (1998). A survey of intonation systems. Hirst, D. and A. Di Cristo (eds), *Intonation Systems. A Survey of Twenty Languages*, 1-44. Cambridge University Press.

Hiz, H. (1978). Difficult Questions. H. Hiz (ed.), *Questions*, 211-226. Dordrecht, Reidel.

Hockey, B.A. (1994). Echo questions, intonation and focus. P. Bosch and J. van der Sandt (eds), *Focus and Natural Language Processing* Vol. II, 93-102: Intonation. IBM Working Papers of the Institute for Logic and Linguistics.

Hoekstra, H. (2000). An algorithm for the assignment of sentence accents in Dutch. H. de Hoop and T. van der Wouden (eds), *Linguistics in the Netherlands 2000*., 105-118. Amsterdam, John Benjamins.

House, J. (1984). Review of C. Gussenhoven (1984), On the Grammar and Semantics of Sentence Accents. *Linguistics* 22, 899-912.

Huddleston, R. (1994). The contrast between interrogatives and questions. *Journal of Linguistics* 30:2, 411-439.

Inkelas, S. and W. R. Leben (1990). Where phonology and phonetics intersect: the case of Hausa intonation. J. Kingston and M. Beckman (eds), *Papers in Laboratory Phonology I*, 35-57. Cambridge University Press.

Iivonen, A. (1998). Intonation in Finnish. Hirst, D. and A. Di Cristo (eds), *Intonation Systems. A Survey of Twenty Languages*, 311-327. Cambridge University Press.

Jackson, S. and Brashers, D.E. (1994). *Random factors in ANOVA*. Thousand Oaks (Calif.), Sage.

Jun, S. A. and M. Oh (1996). A prosodic analysis of three types of wh-phrases in Korean. *Language and Speech* 39 (1), 37-61.

Keijsper, C. E. (1984). Vorm en betekenis in Nederlandse toonhoogtecontouren pt. 1 & 2. *Forum der Letteren* nr. 25, 20-37, 113-126.

Kiefer, F. (1980). Yes-no questions as wh-questions. J. R. Searle, F. Kiefer and M. Bierwisch (eds), *Speech Act Theory and Pragmatics*, 97-119. Dordrecht, Reidel.

Kiefer, F. (1981). Questions and Attitudes. W. Klein and W. Levelt (eds), *Crossing the Boundaries in Linguistics*, 159-176. Dordrecht, Reidel.

Kiefer, F. (1983). Introduction. F. Kiefer (ed.), *Questions and answers*, 1-8. Dordrecht, Reidel.

Kirsner, R. S. and V. J. van Heuven (1996). Boundary tones and the semantics of the Dutch final particles *hè, hoor, zeg* and *joh*. C. Cremers and M. den Dikken (eds), *Linguistics in The Netherlands 1996*, 133-146. Amsterdam, John Benjamins.

Kiss, K. É. (1998). Identificational focus versus information focus. *Language* (74), 245-273.

Kraaijeveld, J. (1997). *Idiosyncrasy in prosody. Speaker and speaker group identification in Dutch using melodic and temporal information.* PhD thesis, University of Nijmegen.

Kretschmer, P. (1938). Der Ursprung des Fragetons & Fragesatzes. Ambrogio Ballini *et al*. (eds), *Scritti in Onore di Alfredo Trombetti*, 27-50. Milano, Ulrico Hoepli.

Ladd, D. R. (1980). *The structure of intonational meaning.* Bloomington, Indiana University Press.

Ladd, D. R. (1981). On Intonational Universals. T. Myers, J. Laver and J. Anderson (eds), *The cognitive representation of speech*, 389-397. Amsterdam, North Holland Publishing Company.

Ladd, D. R. (1983a). Phonological features of intonational peaks. *Language*59, 721-759.

Ladd, D. R. (1983b). Peak Features and Overall Slope. A. Cutler and D. R. Ladd (eds), *Prosody: Models and Measurements*, 39-52. Berlin, Springer.

Ladd, D. R. (1984). Declination: a review and some hypotheses. *Phonology Yearbook* 1, 53-74.

Ladd, D. R. (1987). Review of Bolinger 1986. *Language* 63, 637-643.

Ladd, D. R. (1988). Declination "reset" and the hierarchical organisation of utterances. *Journal of the Acoustical Society of America* 84 (2), 530-545.

Ladd, D. R. (1990). Intonation: Emotion vs. grammar." (Review article on D. Bolinger 1989, *Intonation and its uses,* Stanford University Press). *Language 66, nr. 4,* 806-816.

Ladd, D. R. (1993a). Constraints on the gradient variability of pitch range, or, Pitch level 4 lives! P. Keating (ed.), *Papers in Laboratory Phonology III*, 43-63. Cambridge University Press.

Ladd, D. R. (1993b). On the theoretical status of "the baseline" in modelling intonation. *Language and Speech 36(4),* 435-451.

Ladd, D. R. (1996). *Intonational Phonology.* Cambridge University Press.

Ladd, D. R. (1998). A critical remark on *Intonational Phonology:* response to Nina Grønnum. *Journal of Phonetics* 26, 113-114.

Ladd, D. R. and R. Morton (1997). The perception of intonational emphasis: continuous or categorical?. *Journal of Phonetics 25,* 313-342.

Ladd, D. R. and J. Terken (1995). Modelling intra- and inter-speaker pitch range variation. *International Conference on Phonetic Sciences* 13 (Stockholm), Vol. 2., 386-389.

Ladd, D. R., J. Verhoeven and K. Jacobs (1994). Influence of adjacent pitch accents on each other 's perceived prominence: two contradictory effects. *Journal of Phonetics* 22, 87-99.

Liberman, M. and J. Pierrehumbert (1984). Intonational Invariance under Changes in Pitch Range and Length. M. Aronoff and T. Oehrle (eds), *Language and Sound Structure*, 157-233. Cambridge (Mass.), MIT.

Lieberman, P. (1965). On the Acoustic Basis of the Perception of Intonation by Linguists. *Word* 21, 40-54.

Lindsey, G.A. (1981). Intonation and pragmatics. *Journal of the International Phonetic Association* Vol. 11 nr. 1, 2-21.

Lindsey, G.A. (1985). *Intonation and Interrogation: Tonal Structure and the Expression of a Pragmatic Function in English and Other Languages.* PhD thesis. Los Angeles, University of California

Luksaneeyanawin, S. (1998). Intonation in Thai. D. Hirst and A. Di Cristo (eds), *Intonation Systems*, 376-394. Cambridge University Press.

Lyons, J. (1977). *Semantics*, 2 Vols. Cambridge University Press.

Makarova, V. (1999a). Pitch peak alignment in Russian declaratives, interrogatives, and exclamations. *Proceedings 14th International Conference on Phonetic Sciences*, San Francisco, Vol. 2, 1173-1176.

Makarova, V. (1999b). Perception of intonational contrasts by Japanese and Russian listeners. *Proceedings 14th International Conference on Phonetic Sciences*, San Francisco, Vol. 3, 1945-1948.

Marsi, E.C., P. J. M. Coppen, C. H. M. Gussenhoven and T. C. M. Rietveld (1996). Prosodic and Intonational Domains in Speech Synthesis. J. Van Santen, J. Olive, J. Hirschberg (eds), *Progress in Speech Synthesis*, 477- 493. New York, Springer.

Meeuwissen, L. *Spreekuur of Zwijguur.* PhD dissertation, Nijmegen University.

Mennen, I. (1999). *Second language acquisition of intonation: the case of Dutch near-native speakers of Greek*. PhD thesis, University of Edinburgh.

Mey, J. L. (1993). *Pragmatics.* Oxford, Blackwell.

Moraes, J. A. (1998). Intonation in Brazilian Portuguese. Hirst, D. and A. Di Cristo (eds), *Intonation Systems. A Survey of Twenty Languages*, 179-194. Cambridge University Press.

Mozziconacci, S.J.L. (1998). *Speech variability and emotion: Production and perception.* PhD thesis, Eindhoven University.

Myers, S. (1999). Downdrift and pitch range in Chichewa intonation. *Proceedings 14th International Conference of Phonetic Sciences*, San Francisco, 1981-1984.

Nakajima, S. and J. F. Allen (1993). A Study on Prosody and Discourse Structure in Cooperative Dialogues. *Phonetica* (50), 197-210.

Nolan, F. (1995). The effect of emphasis on declination in English intonation. J. Windsor-Lewis (ed.), *Essays in Honour of Professor J. D. O'Connor*, 241-254. London, Routledge.

Nooteboom, S. G. and J. G. Kruyt (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *Journal of the Acoustical Society of America* 82 (5), 1512-1524.

Ohala, J. (1983). Cross-Language Use of Pitch: An Ethological View. *Phonetica* 40, 1-18.

Ohala, J. (1984). An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice. *Phonetica* 41, 1-16.

Ohala, J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. L. Hinton, J. Nichols and J. Ohala (eds), *Sound symbolisms,* 325-347. Cambridge University Press.

Oppenrieder, W. (1988). Intonatorische Kennzeichnung van Satzmodi. H. Altmann (ed.), *Intonationsforschungen*, 169-205. Tübingen, Niemeyer.

O'Shaughnessy, D. (1979). Linguistic features in fundamental frequency patterns. *Journal of Phonetics* 7, 119-145.

Overdiep, G. S. and G. A. van Es (1949). *Stilistische Grammatica van het Moderne Nederlandsch.* Zwolle, Tjeenk Willink.

Paardekooper, P. C. (1971). *Beknopte ABN-Syntaksis.* Den Bosch, Malmberg.

Parkes, M. B. (1992). *Pause and Effect. An Introduction to the History of Punctuation in the West.* Aldershot (Hants, UK), Scolar Press.

Patterson, D. and D. R. Ladd (1999). 'Pitch range modelling: linguistic dimensions of variation'. *Proceedings 14th International Conference on Phonetic Sciences,* San Francisco, 1169-1172.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation.* PhD thesis, MIT, published 1988 by IULC.

Pierrehumbert, J. B. and M. E. Beckman (1988). *Japanese Tone Structure.* Cambridge (Mass.), MIT.

Pierrehumbert, J. B. and J. Hirschberg (1990). 'The meaning of intonational contours in the interpretation of discourse'. P. R. Cohen, J. Morgan and M. E. Pollack (eds), *Intentions in communication*, 271-311. Cambridge (Mass.), MIT Press.

Pike, K. (1945/1972), *The Intonation of American English.* Ann Arbor, University of Michigan Press.

Post, B. (2000). *Tonal and Phrasal Structures in French Intonation.*. PhD thesis University of Nijmegen, Den Haan, Thesus.

Prieto, P., C. Shih and H. Nibert (1996). 'Pitch downtrend in Spanish'. *Journal of Phonetics* 24, 445-473.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik (1987), *A Grammar of Contemporary English*. Harlow (Essex, UK), Longman.

Refice, M., M. Savino and M. Grice (1997). 'A contribution to the estimation of naturalness in the intonation of Italian spontaneous speech.' *Proceedings ESCA, Eurospeech97*, Rhodes, Greece, 783-786.

Rietveld, T. and R. van Hout (1993). *Statistical Techniques for the Study of Language and Language Behaviour.* Berlin, Mouton de Gruyter.

Rump, H. H. and R. Collier (1996). 'Focus conditions and the Prominence of Pitch-Accented Syllables'. *Language and Speech,* 39 (1), 1-17.

Ryalls, J., G. le Dorze, N. Lever, Lisa Quellet, C. Larfeuil (1993). 'The effects of age and sex on speech intonation and duration for matched statements and questions in French'. *Journal of the Acoustical Society of America* 95 (4), 2274-2276.

Sadock, J. M. (1974). *Toward a linguistic theory of speech acts.* New York, Academic Press.

Sadock, J. and A. Zwicky (1985). 'Speech act distinctions in syntax'. T. Shopen (ed.), *Language Typology and Syntactic Description*, Vol. 1, 155-196. Cambridge University Press.

Scherer, K. R., D. R. Ladd, K. E. A. Silverman (1984). 'Vocal cues to speaker affect: Testing two models'. *Journal of the Acoustical Society of America* 76 (5), 1346-1356.

Schmerling, S. (1976). *Aspects of English sentence stress*. Austin, University of Texas Press.

Searle, J. R. (1976). 'A Classification of Illocutionary Acts'. *Language in Society* 5, 1-23.

Selkirk, E. (1984). *Phonology and syntax: the relation between sound and structure.* Cambridge (Mass.), MIT Press.

Selting, M. (1992). 'Prosody in conversational questions'. *Journal of Pragmatics 17,* 315-345.

Shannon, C. E. and W. Weaver (1949/1964), *The Mathematical Theory of Communication*. Urbana, University of Illinois Press.

Sluijter, A. M. C. and J. M. R. Terken (1993). 'Beyond Sentence Prosody: Paragraph Intonation in Dutch. *Phonetica* (50), 180-188.

Sosa, J. M. (1991). *Fonetica y fonologia de la entonacion del Espanol Hispanoamericano.* PhD thesis, University of Massachusetts.

Sperber, D. and D. Wilson (1986). *Relevance*. Oxford, Basil Blackwell Ltd.

Steedman, M. (1994). 'Remarks on Intonation and "Focus" . P. Bosch and R. van der Sandt (eds), *Focus and Natural Language Processing* Vol. II: Intonation, 185-204. IBM Working Papers of the Institute for Logic and Linguistics.

Strik, H. and L. Boves (1995). Downtrend in F_0 and P_{sb} . *Journal of Phonetics* 23, 203-220.

Studdert-Kennedy, M. and K. Hadding (1973). Auditory and linguistic processes in the perception of intonation contours. *Language and Speech* 16, 293-313.

Stutterheim, C. F. P. (1953). Functie en intonatie. *De Nieuwe Taalgids* XLVI, 129-135.

Swerts, M., D. G. Bouwhuis and R. Collier (1994). A perceptual study of melodic cues to 'finality'. *Journal of the Acoustical Society of America* 96 (4), 2064-2075.

Swerts, M. and R. Geluykens (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech* 37 (1), 21-43.

Taglicht, J. (1982). Intonation and the assessment of information. *Journal of Linguistics* 18, 213-230.

Taylor, S. and R. Wales (1987). Primitive mechanisms of accent perception. *Journal of Phonetics* 15, 235-246.

Terken, J. and D. Hermes (2000). The perception of prosodic prominence. M. Horne (ed.), *Prosody: Theory and Experiment*, 89-127. Dordrecht, Kluwer.

Terken, J. and J. Hirschberg (1994). Deaccentuation of words representing given information: effects of persistence of grammatical function and surface position. *Language and Speech* 37, 125-145.

Terken, J. and G. Nooteboom (1987). Opposite effects of accentuation and deaccentuation on verification latencies for Given and New information. *Language and Cognitive Processes* Vol. 2, Nos. 3/4, 145-163.

't Hart, J. (1998). Intonation in Dutch. D. Hirst and A. Di Cristo (eds), *Intonation Systems*, 96-111. Cambridge University Press.

't Hart, J. and R. Collier (1975). Integrating different levels of intonation analysis. *Journal of Phonetics* 3, 235-255.

Thorsen, N. (1980). A study of the perception of sentence intonation – Evidence from Danish. *Journal of the Acoustical Society of America* 67 (3), 1014-1030.

Thorsen, N. (1985). Intonation and text in Standard Danish. *Journal of the Acoustical Society of America* 77 (3), 1205-1216.

Thorsen, N. (1986). Sentence intonation in textual context-Supplementary data. *Journal of the Acoustical Society of America* 80 (4), 1041-1047.

Tielen, M. J. T. (1992). *Male and female speech. An experimental study of sex-related voice and pronunciation characteristics.* PhD thesis, University of Amsterdam.

Uldall, E. T. (1962). Ambiguity: Question or statement?, or, "Are you asking me or telling me?". *Proceedings 4th International Conference on Phonetic Sciences*,779-783. Mouton & Co., The Hague.

Uldall, E. T. (1972). Dimensions of Meaning in Intonation. D. J. Bolinger (ed.) *Intonation*, 250-259. Hammondsworth, Penguin.

Ultan, R. (1978). Some General Characteristics of Interrogative Systems.J. Greenberg (ed.), *Universals of human language,* Vol. 4: Syntax, 211-248. Stanford University Press.

Vaissière, J. (1983). Language-Independent Prosodic Features. A. Cutler and D. R. Ladd (eds), *Prosody: models and measurements*, 53-66. Heidelberg, Springer.

Van Alphen, J. (1914). De vraagzin. *De Nieuwe Taalgids VIII*, 88-95.

Van Bezooijen, R.A.M.G. (1984). *Characteristics and recognizability of vocal expressions of emotion*. Dordrecht, Foris.

Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech* 38 (3), 253-265.

Van den Berg, B. (1958/1972), *Foniek van het Nederlands.* Den Haag, Goor.

Van den Berg, R., C. Gussenhoven and T. Rietveld (1992), Downstep in Dutch: Implications for a model. G. Docherty and D. R. Ladd (eds), *Laboratory Phonology I*, 335-359. Cambridge University Press.

Van Donzel, M. (1999). *Prosodic aspects of information structure in discourse.*PhD thesis University of Amsterdam. The Hague, Holland Academic Graphics.

Van Donzel, M., F. J. Koopmans-van Beinum and L. C. W. Pols (1998). Speaker strategies in the use of prosodic means in spontaneous discourse in Dutch. *Proceedings ESCA Workshop Sound Patterns of Spontaneous Speech*, Aix-en-Provence, 135-138.

Van Es, G.A. (1932). Syntaxis en dialectstudie: intonatie en syntaxis, 1, 2, 3. *Onze Taaltuin 1,* 87-94, 122-128, 169-175.

Van Es, G.A. (1935). Syntactische functies der intonatie in de volkstaal onzer Noordelijke provinciën. *Handelingen 16e Nederlandsche philologen-congres,* 39-41.

Van Haeringen, C.B. (1962). Vragen over de vraag. *Gramarie: keur uit het werk van zijn hoogleraarstijd*, 288-308. Assen, van Gorcum.

Van Heuven, V. J. and J. Haan (2000a). Phonetic Correlates of Statement versus Question Intonation in Dutch. A. Botinis (ed.), *Intonation. Analysis, Modelling and Technology*, 119-143. Dordrecht, Kluwer Academic Publishers.

Van Heuven, V. J. and J. Haan (2000b). When and how do we hear whether a Dutch speech utterance is a statement or a declarative question?. H. de Hoop and T. van der Wouden (eds), *Linguistics in The Netherlands 2000*, 93-104. Amsterdam, John Benjamins.

Van Heuven, V. J. and J. Haan (2002, to appear). Temporal distribution of interrogativity markers in Dutch: A perceptual study. C. Gussenhoven and N. Warner (eds), *Papers in Laboratory Phonology VII.* Mouton de Gruyter.

Van Heuven, V. J., J. Haan and J. Pacilly (1997). Automatic recognition of sentence type from prosody in Dutch. *Proceedings Eurospeech* 1997 Rhodos, Volume 3, 1483-1486.

Van Heuven, V. J., J. Haan and J. Pacilly (1998). Global and local characteristics of Dutch questions in play-acted and spontaneous speech. D. Duez (ed.), *Proceedings Esca Workshop Sound Patterns of Spontaneous Speech, Production and Perception,* Aix-en-Provence, 139-142.

Van Heuven, V.J., J. Haan, E. J. Van der Torre and Esther Janse (1997). Perceptual identification of sentence type and the time-distribution of prosodic interrogativity. A. Botinis, G. Kouroupetroglou and G. Caryannis (eds), *Proceedings ESCA Workshop on Intonation* 1997, Athens, 317-320.

Van Heuven, V. J. and R. Kirsner (1999). Interaction of Grammatical Form and Intonation: Two Experiments on Dutch Imperatives. R. van Bezooijen and R. Kager (eds), *Linguistics in the Netherlands 1999*, 81-96. Amsterdam, John Benjamins.

Van Wijk, N. (1939). *Phonologie.* 's Gravenhage, Nijhoff.

Volskaya, N. (1999). The influence of the segment context on the realisation of rising F0 patterns in Russian. *Proceedings 14th International Conference of Phonetic Sciences* 1999 San Francisco, Vol. 3, 1985-1987.

Wales, R. and S. Taylor (1987). Intonation cues to questions and statements: how are they perceived?. *Language and Speech* 30-3, 199-211.

Wichmann, A., J. House and T. Rietveld (2000). Discourse Constraints on F0 Peak Timing in English. A. Botinis (ed.), *Intonation. Analysis, Modelling and Technology*, 163-182. Dordrecht, Kluwer Academic Publishers.

Wilson, D. and D. Sperber (1988). Mood and the Analysis of Non-declarative Sentences. J. Dancy, J. Moravcik and C. Taylor (eds), *Human Agency: Language, Duty and Value*, 77-101. Stanford University Press.

Wunderlich, D. (1981). Questions about questions. W. Klein and W. Levelt (eds), *Crossing the Boundaries in Linguistics*, 131-158. Dordrecht, Reidel.

Zwaardemaker, H. and L. P. H. Eijkman (1928). *Leerboek der Phonetiek.* Haarlem, Erven Bohn.

APPENDIX

Experiment information-confirmation questions (chapter 5)

I. Sample set: test utterance in all conditions:

1. Hebben ze mooi weer voorspeld?
2. Ze hebben mooi weer voorspeld?
3. Hebben ze geen mooi weer voorspeld?
4. Ze hebben geen mooi weer voorspeld?
5. Hebben ze dus mooi weer voorspeld?
6. Ze hebben dus mooi weer voorspeld?
7. Hebben ze dus geen mooi weer voorspeld?
8. Ze hebben dus geen mooi weer voorspeld?
9. Ze hebben mooi weer voorspeld hè?
10. Ze hebben geen mooi weer voorspeld hè?
11. Ze hebben dus mooi weer voorspeld hè?
12. Ze hebben dus geen mooi weer voorspeld hè?

II. Instruction form

Geachte proefpersoon,

Hieronder ziet u 36 vraagzinnen. Welk antwoord verwacht de vraagsteller te krijgen?

1. **Het antwoord is voor de vraagsteller volledig onvoorspelbaar, het kan even goed 'ja' als 'nee' zijn.**
   Kruis in dat geval *maximaal onvoorspelbaar* aan.

OF

2. **Het antwoord is voor de vraagsteller min of meer voorspelbaar, d.w.z. hij verwacht al bij voorbaat een bepaald antwoord; hij stelt de vraag eigenlijk vooral om in die verwachting bevestigd te worden.**
   Als het verwachte antwoord 'ja' is, kruis dan *voorspelbaar 'ja'* aan. Verwacht de vraagsteller 'nee', kruis dan *voorspelbaar 'nee'* aan. Geef daaronder een percentage dat de mate van voorspelbaarheid uitdrukt: is dat 100%, of ligt het lager? Omcirkel in dat laatste geval een passend percentage, of zet een kruisje tussen twee aangegeven percentages in.

Voorbeeld:

vraag: **Er zijn dus geen kaarten meer te krijgen?**

antwoord:

maximaal onvoorspelbaar ❍ *voorspelbaar 'ja'* ❍ *voorspelbaar 'nee'*

- 100%
- minder, nl. 80%…60%…
  40%…20%…