# 2 Modeling Danish Intonation

Nina Grønnum

## 2.1 Introduction

Danish belongs to the Nordic (North Germanic) group of Germanic languages. It is the national language of Denmark, and according to Lützen (2010), it is the mother tongue of 92 to 93 percent of Denmark's 5.82 million inhabitants (as of 2020). It is written with the Latin alphabet, extended with *æ*, *ø*, and *å*. Grønnum (2000)[1] is a very brief, and Grønnum (2007) a more comprehensive, modern introduction to its phonetics. The phonology is exhaustively described in Basbøll (2005).

There are three prosodic components in the Danish sound system: stress, stød, and intonation. If vowel length is considered a syllable prosody, it will make a fourth. The phonology of stress and stød is accounted for in Basbøll (2005), and Grønnum, Vazquez-Larruscaín, and Basbøll (2013) contains a more recent discussion specifically about the abstract representation of stød.

This chapter is about intonation only. It is a summary of acoustic analyses and a few perceptual experiments carried out from the mid-1970s through the 1990s (Thorsen 1978 through 1988; Grønnum 1990 through 1998). It presents a model of the acoustic output from Copenhagen Danish speakers' production in distinct, formal speech, of various types of utterances, as well as short texts. Comparison with Central Swedish will highlight the characteristic properties of Danish intonation.

## 2.2 Previous Descriptions

The first linguist to concern himself with Danish intonation was the brilliant J. P. Høysgaard (1747, 1769). According to him, stressed syllables are associated with high tone and unstressed syllables with low tone. He denies intonation any real linguistic status, because words can be comprehended even when whispered. Tone helps only to make spoken language clearly audible.

In the literature of the nineteenth and first half of the twentieth centuries, there is likewise an almost general consensus that stressed syllables have high pitch (Jensen 1833; Sneedorff-Birch 1835; Jerndorff 1897, 1911; Forchhammer 1903; Oxenvad 1937; Arnholtz 1939–1940, 1942, 1954). However, Andersen (1954) notes that in North Zealand and in Copenhagen, stressed syllables are spoken on a lower pitch. To Arnholtz, this is characteristic of the vulgar speech of the lower classes in Copenhagen, where a loathsome change in the stress~pitch relation is taking place: stressed syllables with low or falling pitch—traditionally called *omlæggertone* (approximately "reverse tone"). As recently as 1975, Brink and Lund repeat Arnholtz's observation (albeit in less condemning terms) and state (608–610) that Low Copenhagen often has a falling pitch

movement where High Copenhagen has a rising pitch, and vice versa, Low Copenhagen rises where High Copenhagen falls. Jerndorff (1897) observes that the degree of relative prominence on each stressed syllable correlates directly with its pitch height. The more weight a word is given, the higher the pitch of its stressed syllable. But he added that highlighting of a stressed syllable can be achieved by lowering the preceding unstressed syllable as well.

There are also scattered remarks about sentence modality and tone. Statements are falling and/or end low, and questions are rising and/or end high—or at least they are not falling (Bloch 1805; Mikkelsen 1894; Jerndorff 1897; Forchhammer 1903; Oxenvad 1937, Hjelmslev 1954). In Hjelmslev's (1951) terminology, there are two modulations: one falling and one nonfalling. He notes specifically that these modulations cannot be partitioned into smaller structural domains. Jerndorff (1911) takes issue with the rule for declamation (in schools, lecture halls, churches, and on the stage) that questions must end in a final rise through the unstressed syllables. Other languages may do so, but Danish does not, and he ascribes the role of *carrier of intonation* to the stressed syllables only. Arnholtz (1954) likewise states that in the Danish standard language, it seems that the stressed syllables carry the speech melody, the unstressed syllables being subordinate to the pitch pattern established by the stressed ones.

In unpublished lecture notes from seminars in 1946 and 1947 about Danish speech rhythm and melody, Eli Fischer-Jørgensen makes the same assertion about stressed syllables being the carriers of the speech melody. She discusses at some length the possible criteria for delimiting Danish utterances into stress groups, particularly the problem of assigning the boundaries between them.

Until the 1970s, the only comprehensive accounts of Danish prosody were Otto Jespersen's *Fonetik* (1897–1899, chapters 25 and 26), excerpts of which are found in his *Modersmålets Fonetik* (1934, chapters 13–16), and Alf Bo's monograph *Tonegangen i Dansk Rigsmål* (1933). Jespersen (1897–1899) talks about the law of termination: a low tone or fall in pitch is employed when we are through and want to finish, while a rise to a higher pitch indicates the nonfinished, nonterminal. Bo establishes rules of intonation mainly by impressionistic observation of his own speech. Questions have rising intonation, but the magnitude of the rise in questions varies. Questions with no lexical or syntactic information about their interrogative status are the most rising; questions with word order inversion or interrogative particle less so and may approach declaratives.

Thus, the general picture that emerges from the Danish literature is one where stressed syllables are normally associated with relatively higher pitch than unstressed ones; intonation is carried by the stressed syllables; statements are associated with falling intonation and questions with rising intonation. It is not quite clear whether rises and falls are situated at the end of the utterance or whether they characterize the entire succession of stressed syllables. There are hints that intonation contour trends may be global, especially in some of the transcriptions, but there are also rather clear statements to the effect that only the last stressed syllable is affected.

### 2.3 Preliminaries

The ultimate goal of prosodic theory or, in this case, intonation theory, would be a representation that in a psychologically real way faithfully models every aspect of speakers' and listeners' processing of intonation, including syllable and word tones, in spoken language in general. That is obviously not yet within reach. Settling for a more modest goal, intonation can be modeled from one of several specific vantage points:

as speakers' production or speakers' perception; or as input to synthetic speech, or as patterns for speech recognition—procedures that do not necessarily parallel human processes. Furthermore, speaker and listener processing may be modeled at different levels of abstraction. The first approximation, described in figure 2.10, models two components and their interaction in the production of short utterances in the formal speech of speakers of Copenhagen Danish, leaving out phenomena that typically are not under the speaker's voluntary control, that is, *microprosody*. The more elaborate edition of the model, in figure 2.13, expands the structural domain to longer and syntactically more complex utterances and short texts.

Ideally, a model of the production of Danish intonation is a representation that mediates between a stage upstream in the human production of linguistic utterances, where all the morphosyntactic and lexical information has been supplied and stress assigned to the appropriate syllables, and a stage downstream where significant intonational information is inserted. From this latter stage the phonetic fundamental frequency implementation can be derived by rule and converted to neural commands to drive the peripheral physiological production system or, alternatively, converted to commands to drive a speech synthesizer. Over the last couple of decades, it has become customary to call a representation that in this manner mediates between a less and a more concrete stage in the production of intonation *phonological*. Ladd (2008) contains a lucid discussion of the pros and cons of such a concept. However, if you want to avoid the obvious—but obviously false—analogy to segmental phonology and its sequentially ordered and minimally contrastive units, then *symbolic representation* is perhaps a better concept.

## 2.4 Danish Intonation

It is customary to let "intonation" designate the linguistically relevant but nonlexical aspects of the *course of fundamental frequency* (or its perceptual correlate, the *pitch course*) through spoken utterances (see, e.g., Ladd 1996). Nonlexical leaves out syllable tones and word accents, but it also excludes the contribution to fundamental frequency (henceforth $F_0$) that comes from stress. However, the stress~$F_0$ relation is a salient property of Danish speech melody. Therefore, as long as the subject is Danish, which has neither syllable tones nor tonal word accents, let intonation designate the linguistically relevant aspects of $F_0$ in spoken utterances, including $F_0$ patterns associated with stress. "Linguistically relevant" excludes microprosodic effects, that is, the contribution from intrinsic $F_0$ properties of segments and perturbations at segment boundaries. I take the core linguistic function of intonation to be to signal (i) prominence relations, (ii) prosodic boundaries between phrases and utterances, and (iii) utterance modality.

### 2.4.1 Empirical Investigations

Empirical studies of Danish intonation were long in coming, even well after the appearance of suitable instrumentation for analysis. Intonation was considered difficult to approach, for reasons stated by Poul Andersen (1954, 310) [in my near verbatim translation]: "To give a general description of tonal, stress, and speech rate variations in the sentence is made difficult by the fact that these phenomena vary with the speaker's mood and emotional state, and it is very difficult to decide if we are dealing with, e.g., a neutral question or a surprised question, and, strictly speaking, there really is no such thing as a neutral intonation, since all speech is connected with some emotional state." In other words, $F_0$ codes several types of information simultaneously—for example, the linguistic function, the affective state of the speaker, the speaker's age, gender, personal

identity, and membership of regional and social groups. The difficulty inherent in teasing out the contribution from each of these various factors likely appeared prohibitive.

By the mid-1970s, however, there was a growing demand for descriptions of intonation that could be applied in the teaching of Danish as a foreign language and in teaching the hearing impaired. In spite of Andersen's apprehensions, I set out to discover the linguistically relevant properties in the "tonal variations in the sentence" and provide a description of "neutral intonation." Andersen is right that, strictly speaking, there is no such thing as a neutral utterance or neutral intonation. Every utterance is produced in some context, and the context is one factor in the manner of its production. But we can operationally define two helpful concepts: *unmarked* intonation and *typical* intonation. By definition, subject-verb (SV) declarative sentences are syntactically unmarked in Danish. By convention, the intonation contour that accompanies such an SV declarative when it is used conventionally is unmarked. Typical intonation, on the other hand, is one that accompanies any given sentence type when used conventionally. Thus, for instance, the intonation associated with a conventionally used yes/no question will be different from the unmarked intonation of conventional declaratives; hence it is marked, but that is typical.

### 2.4.2 Methodology

The only way to approach the relation between stress and $F_0$ in typical intonation contours, keeping everything else constant and similarly, to lay bare typical intonation contours, was to have speakers read material designed for the purpose—in this case, short sentences embedded in small dialogues, some of which were built around nonce words with systematically different word stress locations. Details can be found in Thorsen (1978), and suffice it here to say that in this first experiment, as well as every succeeding one, there were always at least four speakers and at least six repetitions of every target utterance by each speaker in randomized scripted material. Note that everything I say about Danish as well as Swedish intonation in the following pertains to utterances conventionally used, recorded in a formal and distinct style of speech.

### 2.4.3 Paucity of Prosodic Components in Danish

Almost every significant fact about Danish intonation can be deduced from the declarative utterance in figure 2.1, with its unmarked and typical declarative intonation

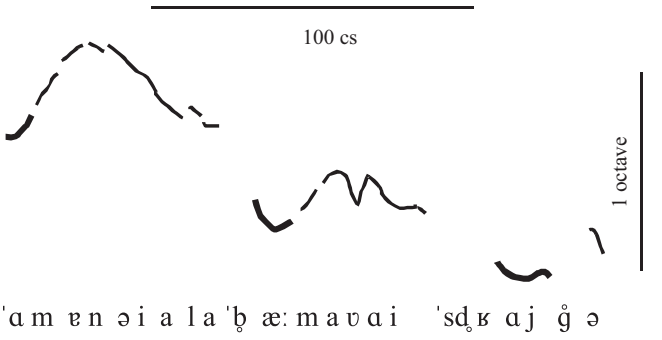


**Figure 2.1**
Fundamental frequency tracing of a female speaker's utterance: *Ammerne i Alabama var i strejke*. The stressed vowels are in heavier lines.

contour: [ˈɑmɐnə i alaˈb̥æːma ʊɑ i ˈsd̥ʁɑjg̊ə] *Ammerne i Alabama var i strejke* (The wet-nurses in Alabama were on strike), composed so as to minimize confounding influence from differences in intrinsic vowel $F_0$ and perturbations at segment boundaries. (This is merely a sample utterance for illustration, backed up by the results of empirical investigations.)

- If the course of $F_0$ is any indication, the six words in this utterance form three prosodic groups: {Ammerne i Ala}{bama var i}{strejke}. Each is characterized by a rise in $F_0$ from the stressed to the first posttonic syllable, a rise whose magnitude decreases from start to finish, followed by a more or less extensive fall through succeeding posttonics, if any. (The small dip halfway through the second pattern is due to the intrinsically lower $F_0$ of the [ʊ].) The extent of the $F_0$ fall varies in accordance with the number of posttonic syllables, and with only one posttonic syllable (as in *strejke*), the fall is absent.
- There is no special and elaborate $F_0$ movement at the end of the utterance.
- There is no pronounced fall to low at the very end.
- The stressed diphthong of *strejke* is not visibly or perceptibly lengthened.

In similar circumstances, a Central Swedish utterance—in addition to the tonal patterns associated with the word accents—would have exhibited a prominent final rising sentence accent, a clear final low boundary tone, and final lengthening—all of which are absent here.

### 2.4.4 The Prosodic Stress Group and Its $F_0$ Pattern

$F_0$ is the most explicit among the acoustic cues to stress. The onset of any stressed vowel coincides with the onset of a recurrent melodic pattern that extends over all succeeding unstressed syllables, irrespective of their morphological or syntactic affiliation, cutting across word boundaries, until the onset of the next stressed vowel. Such a sequence of syllables forms a prosodic stress group, with an associated $F_0$ pattern. This is the stress group that Fischer-Jørgensen intuited in 1946–1947, but whose boundaries she hesitated to determine. Its $F_0$ pattern, with the low stressed syllable relative to the first posttonic, is the one that Arnholtz (1939–1940) found vulgar and loathsome. It is now the norm in Copenhagen Danish and no longer confined to the "lower classes," as Fischer-Jørgensen (1983) also noted in her acoustic analysis of a Copenhagen speaker. The pattern is subject to truncation/extension (rather than compression/expansion), as evidenced in figure 2.1. The shorter or longer the stress group, the less or more extensive is its $F_0$ pattern. It is also sensitive to location: rises (and falls) are more extensive early in the utterance than later, and mostly also slightly more extensive on less declining intonation contours, everything else being equal.

### 2.4.5 Equal and Unequal Prominence

In Standard Danish, in pragmatically and emotionally neutral speech, no one stressed syllable is more prominent than any other. Such prominence is not evident in acoustic registrations (in duration and/or the $F_0$ contour) nor is it present auditorily. There is nothing incomplete about such prosodically neutral utterances, and although they may not be very frequent in spontaneous speech, they certainly do occur, they are not unnatural, and they are very easy to elicit from speakers in a reading situation. This is a completely uncontroversial fact among Danish scholars (Andersen 1954; Basbøll 2005; Heger 1981; Rischel 1980, 1983).

**2.4.5.1 No default sentence accent** Merely looking at the tracing in figure 2.1, you might think that *Ammerne*, with its more ample $F_0$ pattern, sounded more prominent than *strejke*. That is not the case. The three stressed words are equally prominent perceptually. Nor do I expect anyone to read a final sentence accent into the utterance in figure 2.1. A comparison with Swedish is illustrative.

Stockholm and Copenhagen speakers recorded the same utterance in isolation, and in two different dialogue contexts: *Torbens syster heter Kamma/Torbens søster hedder Kamma* (Torben's sister is called Kamma). One context inquired about Torben's sister's name, to elicit focus on *Kamma*; the other asked who had a sister called *Kamma*, to elicit focus on *Torbens* and reduced prominence (deaccentuation) on *Kamma*. The six $F_0$ contours in figure 2.2 depict the last word, *Kamma*, in one recording of each of these six utterances. They are clearly different among themselves in Swedish. In the isolated utterance (left), *Kamma* exhibits a word accent II fall from high, a default sentence accent rise, and a (not very explicit) boundary low, as specified by Bruce (1977). When focused (middle), the sentence accent rise is higher and the fall to low deeper. When the focus is earlier in the utterance (right), the sentence accent rise is absent. The three words in the Danish utterances exhibit nothing comparable and are hardly distinguishable.

**2.4.5.2 Focus by reduction** The isolated Danish utterance, *Torbens søster hedder Kamma*, and the answer that focuses on *Kamma*, from which the tracings in figure 2.2 are drawn, are not distinguishable perceptually when played back out of context, which is not surprising when you look at the top and middle panels in figure 2.3. But that is specific to final words where there is no succeeding stress group to reduce, which is how focus is achieved in this speech style. Thus, the answer to whose sister is called *Kamma*, which focuses on *Torbens*, is clearly distinguishable from the isolated utterance, due not so much to increased $F_0$ on *Tor-*, but rather to the reduced prominence on succeeding *søster*, via deletion of the $F_0$ rise from *søs-* to *-ter* (bottom panel in figure 2.3).

**2.4.5.3 Emphasis for contrast by reduction** With a greater degree of relative prominence, the prominent word itself may be affected, that is, be boosted by a somewhat

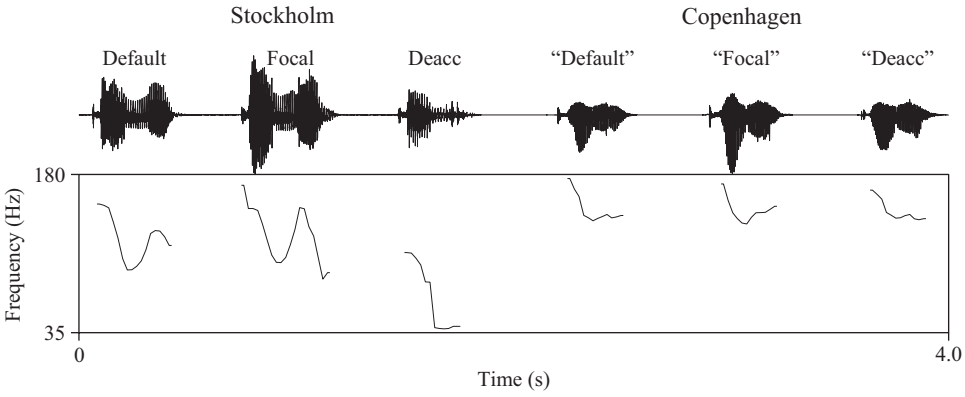


**Figure 2.2**
Fundamental frequency tracings of the last word, *Kamma*, in three short utterances in Stockholm Swedish (male speaker, left) and Copenhagen Danish (female speaker, right).

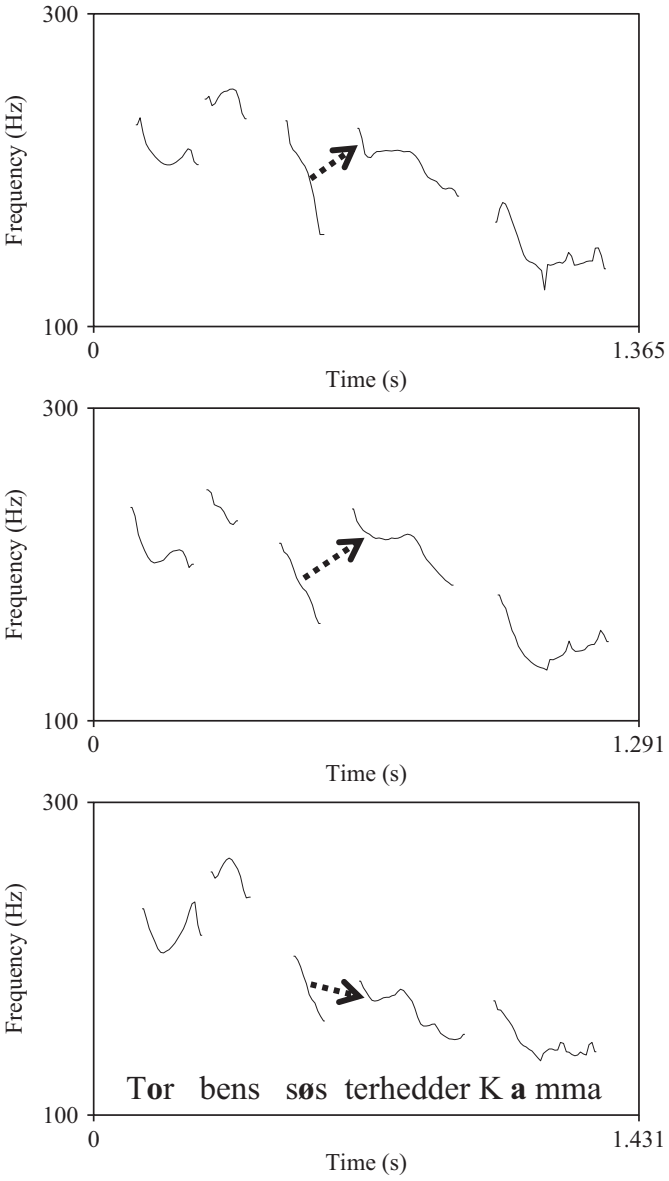


**Figure 2.3**
Three versions of *Torbens søster hedder Kamma* (female speaker): in isolation (top), with focus on *Kamma* (middle), and on *Torbens* (bottom).

increased $F_0$ on the stressed syllable, as well as a higher rise to the posttonic syllable. But equally significant is the reduction on preceding as well as succeeding stress groups, as illustrated in figure 2.4. It depicts an utterance, *Der går mange busser fra Tiflis* (There are many buses out of Tiflis), in isolation and as a response to a question that inquired whether there were in fact buses out of Tiflis or if one had to go by train. The $F_0$ rises on *mange* and *Tiflis* are deleted, and the rise from *bus-* is higher in the emphasized version on the right. Jerndorff's (1897) observation, that the degree of relative prominence on

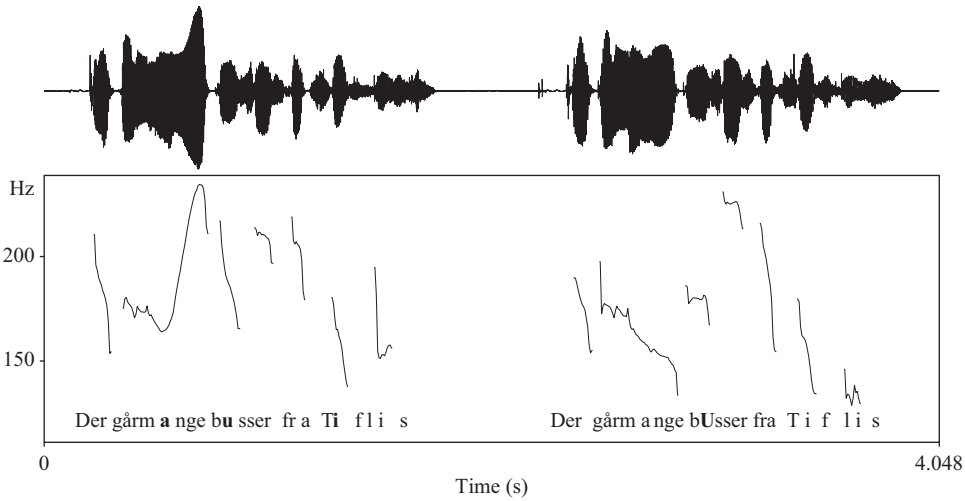


**Figure 2.4**
Fundamental frequency tracings of two versions by a female speaker of *Der går mange busser fra Tiflis* uttered in isolation (left) and with emphasis for contrast on *busser* (right).

a stressed syllable need not be achieved by increasing its pitch but can be accomplished also by lowering the preceding unstressed syllable, was not wide of the mark.

### 2.4.6 No Final Boundary Tone

The absence of a low, final boundary tone in Copenhagen is very apparent, also in comparison with Danish as spoken on Bornholm, a Danish island just south of Sweden in the Baltic Sea. Figure 2.5 depicts $F_0$ tracings of a long declarative utterance: *Kofoed og Thorsen skal med rutebilen fra Gudhjem til Snogebæk klokken fire på tirsdag* (Bornholm, top panel) and *Kofoed og Thorsen skal med bussen fra Fuglebjerg til Sorø klokken fire på tirsdag* (Copenhagen, bottom panel) (Kofoed and Thorsen are taking the bus from . . . to . . . at four o'clock on Tuesday). For ease of comparison, the frequency scale spans the same interval in both panels.

### 2.4.7 No Final Lengthening

Final syllables generally (but not invariably) exhibit a modest increased duration in Copenhagen Danish, everything else being equal. I used to think of that as final lengthening. However, when I compared it with Central Swedish, I had occasion to reconsider. Figure 2.6 depicts *Kamma* from initial and final position, respectively, in Swedish *Kamma kommer från Svaneke* and *Torbens syster heter Kamma* (left) and Danish *Kamma stammer fra Næstved* and *Torbens søster hedder Kamma* (right) (Kamma comes from Svaneke/Næstved; Torben's sister is called Kamma). The Swedish *Kamma* is 33 percent longer in final than in initial position, against a meager and nonperceptible 7 percent lengthening in Copenhagen. No doubt final lengthening is an independent factor in Swedish, but in Danish, it may be merely the result of a general articulatory relaxation toward the end of phrases and utterances, and not an autonomous, independently speaker-controlled feature. I would argue similarly about the increased phrase final duration that Tøndering (2008) found in the nonscripted monologues in the Danish Phonetically Annotated Spontaneous Speech corpus (DanPASS; see Grønnum 2009). It is unquestionably an acoustic reality, although *not* ubiquitous, but it is nowhere near

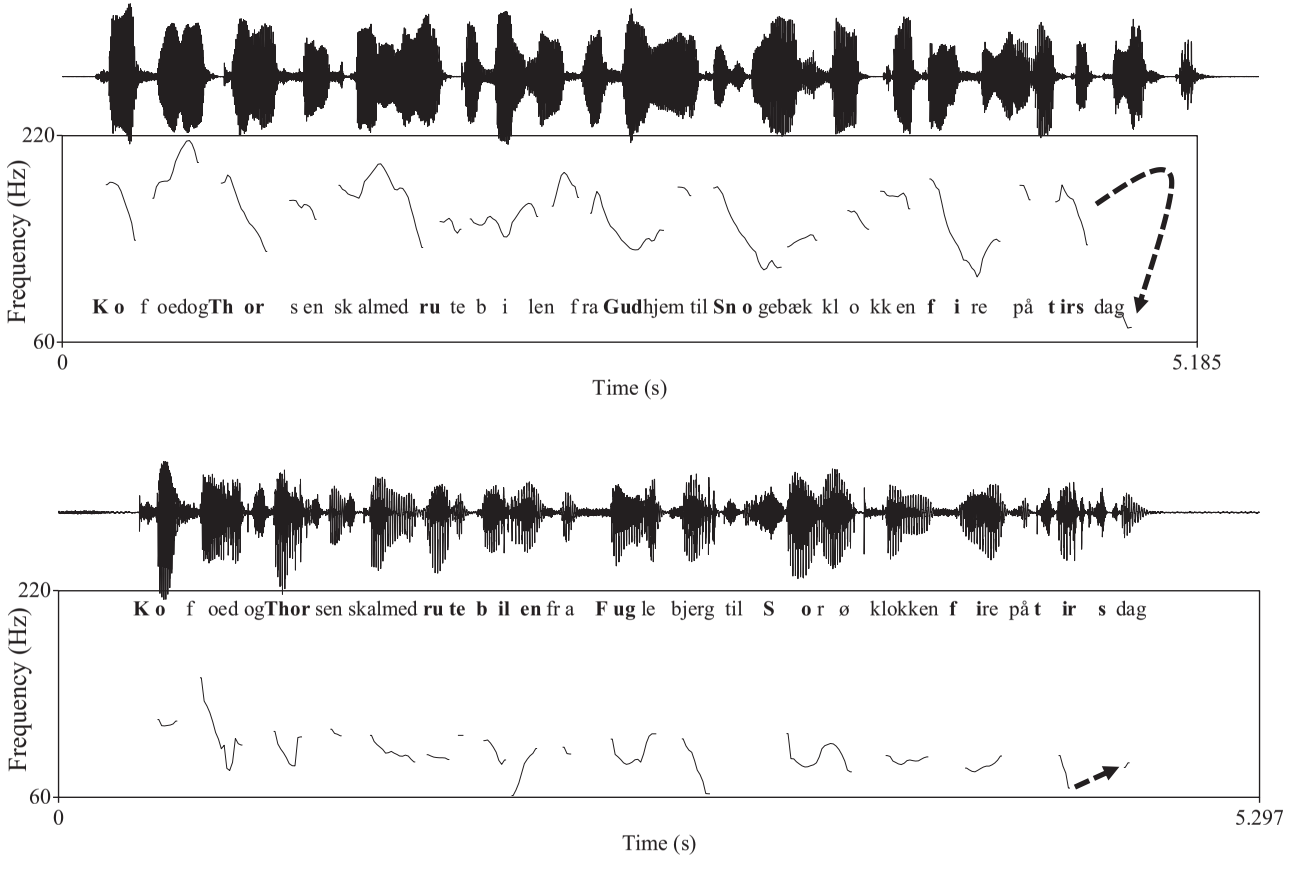


**Figure 2.5**
Fundamental frequency tracings of a long, terminal declarative utterance by male speakers in Bornholm (top) and Copenhagen (bottom).

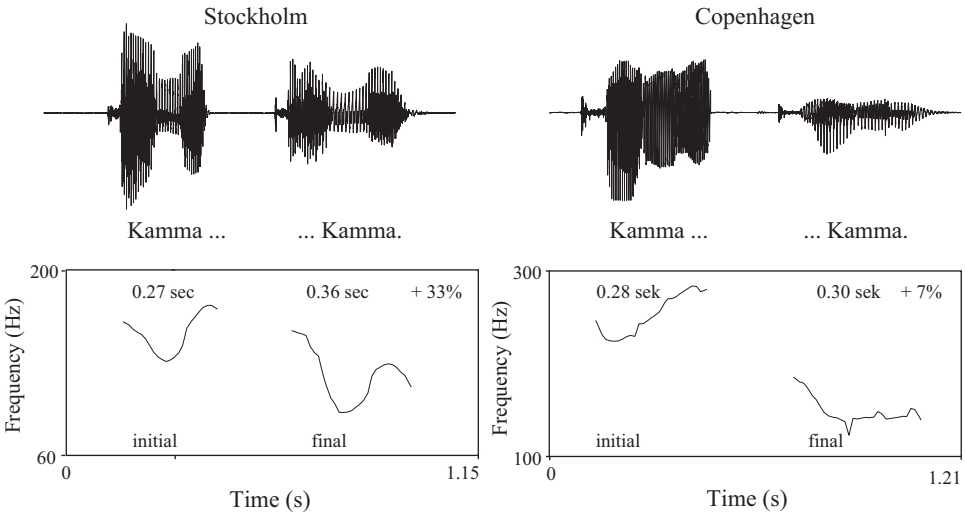


**Figure 2.6**
Fundamental frequency contours of the word *Kamma*, extracted from utterance initial and final position, respectively, in Swedish (left trace, male speaker) and Danish (right trace, female speaker) utterances.

as perceptually salient as in Central Swedish. In other words, final lengthening is not a component in the abstract, symbolic representation of Copenhagen Danish prosody.

### 2.4.8 No Final Rise or "High" in Questions

Compare Danish with Swedish once again. Among my recordings are the following utterances, produced in isolation: *Kamma kommer från Svaneke* (Stockholm; Kamma comes from Svaneke) and *Anders og Kamma skal til Fakse* (Copenhagen; Anders and Kamma are going to Fakse). I have small dialogues as well: *Hur långt är det från Sandvik till Svaneke?–Till Svaneke?–Det är tre mil* (Stockholm; How far is it from Sandvik to Svaneke?–To Svaneke?–It is three miles), and *Hvor langt er der fra Næstved til Fakse?–Til Fakse?–Der er cirka 30 kilometer.* (Copenhagen; How far is it from Næstved to Fakse?–To Fakse?–It is about 30 kilometers).

Figure 2.7 depicts $F_0$ tracings of the last word in the Stockholm and Copenhagen renderings, from the declarative (leftmost of three), the *wh*-question (middle), and the question that is not syntactically marked as such (an intonation question; rightmost of three). As expected, Stockholm exhibits the word accent I low stressed syllable, the sentence accent rise, and the boundary low, where the distinguishing feature is the magnitude of the sentence accent rise, which is somewhat higher in the *wh*-question and very much higher in the intonation question than in the statement. It is not that statements fall or end low in the Swedish utterances and questions rise or end high. Statements and questions alike end with a low boundary tone, but the preceding rise and sentence accent peak are significantly higher in questions, and much more so when there is nothing but the intonation to carry the interrogative modality. In the three Danish words, it is not apparent that the rise itself is significantly higher from *Fak-* to *-se* in either of the two questions, but the whole word is lifted way up in the range in the intonation question. So perhaps there is some truth to an assertion that questions—at least when devoid of lexical or syntactic cues to their modality—end

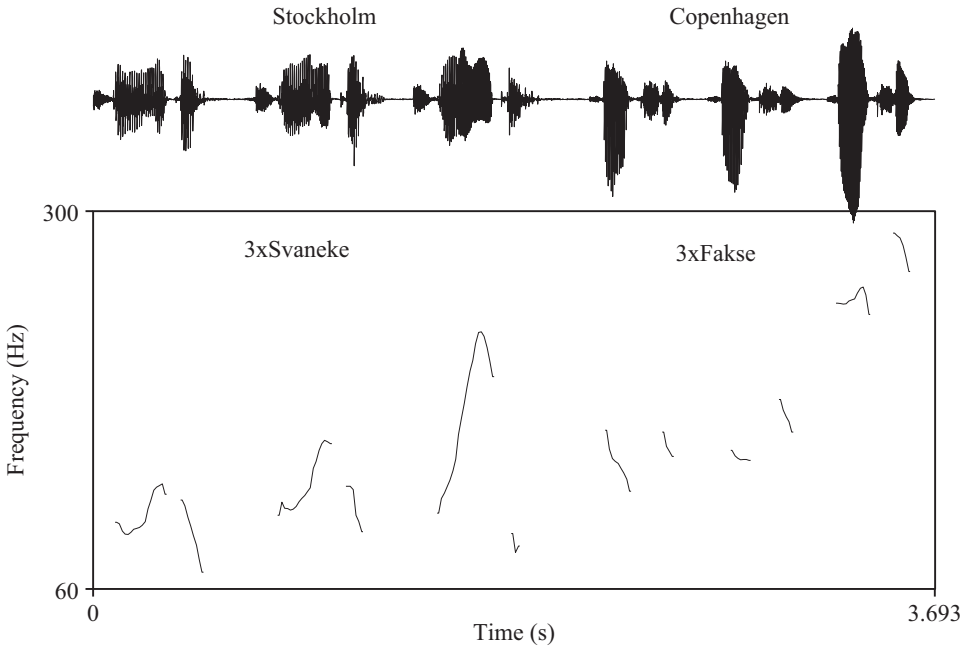


**Figure 2.7**
Fundamental frequency tracings of the last word (*Svaneke* and *Fakse*, respectively) in a statement and two questions in Stockholm (male speaker, left) and Copenhagen (female speaker, right).

high. Such a claim, however, misses the point. The pertinent feature becomes apparent when we look at a complete sentence, as in figure 2.8.

Compare the same sentence, *Der går mange busser fra Tiflis* (There are many buses out of Tiflis), uttered as a terminal declarative in isolation and as a question. The question was the third and last utterance in this little exchange: *Hvornår skal vi afsted? Hvad står der i køreplanen? Der går mange busser fra Tiflis?* (When are we leaving? What does the timetable say? There are many buses out of Tiflis?). It depends on the intonation to be perceived as a request for confirmation. What distinguishes this, the most marked of intonation contours, from the unmarked, declarative intonation, is not the final $F_0$ movement but rather the overall nondeclining contour, suggested by the broken lines. Because the contour is not actually rising, *Tiflis* in the intonation question on the right is not higher than the preceding *busser*. In this light, the rightmost *Fakse* in figure 2.7 should be seen not as high per se but as the termination of a (virtual) nondeclining contour.

There is more to be said about intonation and modality, but that is more easily done after introduction of the initial version of the model presented in section 2.4.10 and figure 2.10.

### 2.4.9 The Stressed Syllables Carry the Intonation Contour

Jerndorff's (1911) and Fischer-Jørgensen's (1946–1947) impression that the stressed syllables carry the intonation is entirely justified. Even more to the point is Arnholtz's (1954) added observation: that the unstressed syllables are subordinate to the pattern established by the stressed ones. It is evident in figure 2.1 that $F_0$ patterns are sensitive to location and diminish through the utterance. Figure 2.9 illustrates how the scaling of

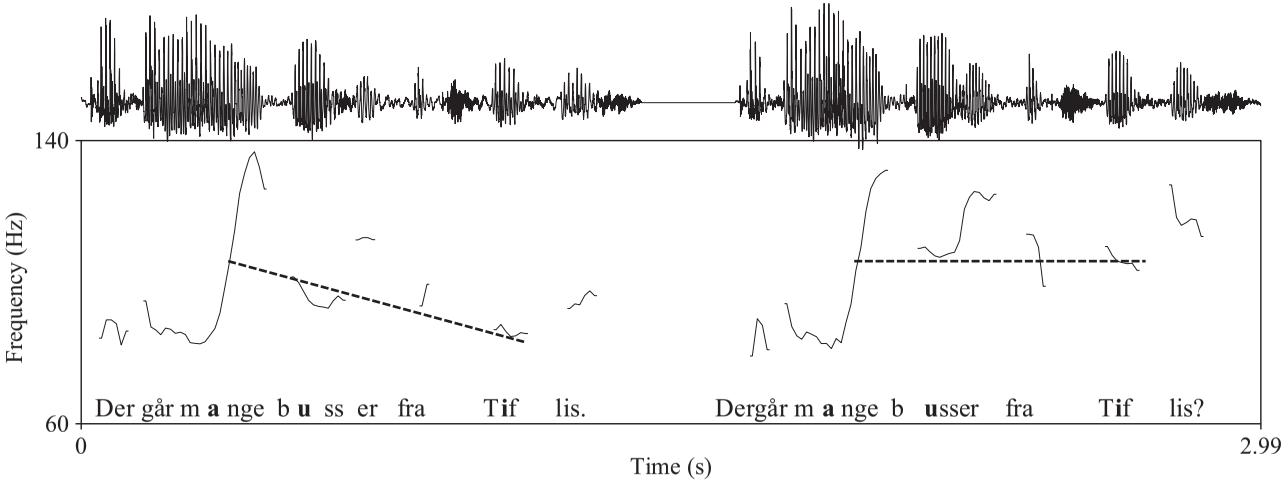


**Figure 2.8**
Fundamental frequency tracings of the same sentence produced as a terminal declarative (left) and a question (right), male speaker. Dotted lines are drawn through the three stressed vowels.

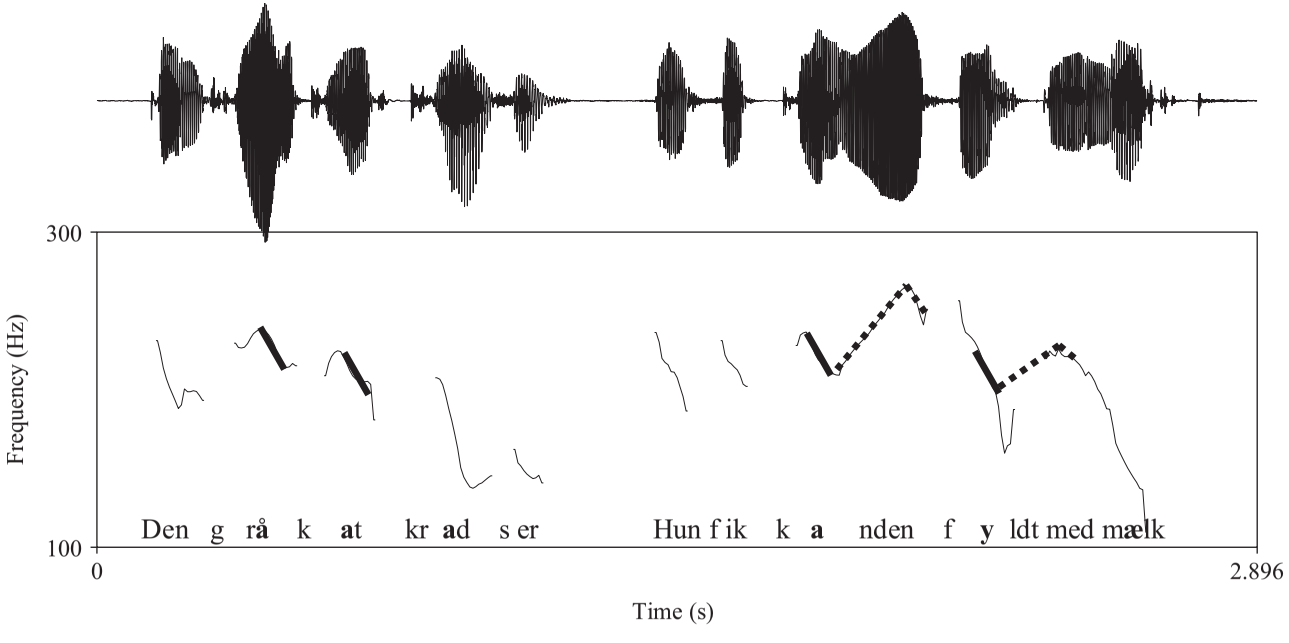


**Figure 2.9**
Fundamental frequency tracings of *Den grå kat kradser* (left) and *Hun fik kanden fyldt med mælk* (right), female speaker.

the stressed syllables is unaffected by the presence, or not, of unstressed syllables in the stress group. On the left is an utterance, *Den grå kat kradser* (The gray cat scratches), where the first two stressed syllables are without posttonics. On the right is *Hun fik kanden fyldt med mælk* (She had the jug filled with milk). *Grå* and *kan-* appear on approximately the same level, as do *kat* and *fyldt*, whether or not they are succeeded by rises to a posttonic.

$F_0$ patterns, specifically the magnitude of their rise, are also sensitive to the intonation contour on which they ride, generally rising slightly higher on marked contours, everything else being equal. If we disregard contextually induced focus or emphasis, we are left with $F_0$ patterns in prosodic stress groups, whose manifestation is entirely predictable from the wider context, early or late, higher or lower on the (more or less declining) intonation contour. It is hard to say to what extent this variation is voluntary and directly speaker controlled or should be ascribed to general speech production principles, which reduce articulatory explicitness through time or in unmarked versus marked contexts. But irrespective of the output control mechanism and even though $F_0$ patterns are an integral part of a model of speaker performance, they are not part of the abstract underlying representation, properly speaking. Stress belongs there, but its phonetic manifestation does not. Being contextually determined, however, does not deprive the predictable $F_0$ pattern variation of any perceptual relevance. It becomes the norm against which noneven distribution of prominence is evaluated.

Similar views about intonation as a layered structure are to be found in Bruce (1977), Fujisaki, Hirose, and Ohta (1979), Gårding (1979), 't Hart and Collier (1979), Möbius (1993), Öhman (1968), and Vaissière (1983). Note, however, that the Danish model presented in section 2.4.10 and figure 2.10 is not simply an additive "overlay" model. The intonation contour determines the execution of the prosodic phrase contour, which in its turn sets the scale for the prosodic stress group pattern, that is, *superposition* and *subordination* are equally important aspects of the model. (The full documentation for the descriptive adequacy of the model does not find room here, but it can be found in previous work, all of it referenced and most of it summarized in Grønnum 1992.)

### 2.4.10 Modeling Short Utterances

Having done away with most of those parameters that are integral parts of other languages (and thus Standard Central Swedish)—sentence accents, boundary tones, and local final cues to modality, as well as final lengthening—we are left with the manifestation of stress and the overall, global intonation contour (see figure 2.10).

**2.4.10.1 Utterance intonation contours** The central assumption, borne out by prior acoustic analyses, is that the stress group patterns have no autonomous status other than as the manifestation of stress. They are superposed on and subordinate to the intonation contour where they occur, rises being higher early than late and higher on less falling contours, everything else being equal. The stressed syllables are frequency scaled independent of the presence or not of unstressed syllables. Furthermore, $F_0$ patterns are truncated rather than compressed when time is short (see figures 2.1 and 2.9), which means that there is no upward deflection of $F_0$ in a stress group without posttonics. And in a string of, say, four stressed syllables, as in, for instance, *Lis så Pers film* (Lis watched Per's film), the stressed syllables are frequency scaled in the same way as in *Lissi kendte Peters tante* (Lissi knew Peter's aunt), but their timing will be different. Carry this observation over to the model in figure 2.10. If you lift out one or more dots from a stress group, you should cut back the thin full line accordingly.

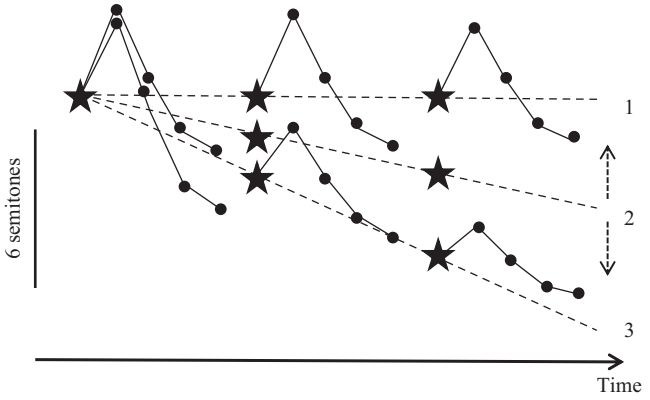


**Figure 2.10**
A model for the course of fundamental frequency in short utterances in Standard Copenhagen Danish. Stars represent stressed syllables and dots unstressed ones. The solid lines depict the $F_0$ patterns associated with prosodic stress groups, and dashed lines denote the intonation contours proper: 1, syntactically unmarked questions; 2, interrogatives with word order inversion and/or interrogative particle (variable), and nonterminal clauses; 3, terminal declaratives.

Intonation contours vary between most steeply declining—in declarative sentences used conventionally (line 3 in figure 2.10), and level—in utterances that are not marked lexically or syntactically for their interrogative function (line 1 in figure 2.10). Other questions and nonfinal clauses fill in the intermediate space, with a clear tendency for a trade-off between lexical/syntactic markers and intonation contour slope (line 2 in figure 2.10). The steepest contours typically span four to six semitones. The choice of contour slope for a given utterance is determined by syntactic and, not least, pragmatic factors, in accordance with principles of markedness and typicality, as I have suggested. Syntactically unmarked declarative sentences used conventionally are associated with unmarked intonation contours, and they are the most steeply declining of all. That makes any less falling contour marked. Typical intonation is the contour that accompanies any given sentence type when it is used conventionally. Thus, a conventional yes/no question will have a slope somewhere between unmarked (most steeply declining) and level, which means it is marked but typical. Any deviation from the typical intonation will be contrary to conventional use and have implications for illocutionary force. Thus, the most strongly marked intonation contour, line 1 in figure 2.10, will, by virtue of its nontypicality, turn a declarative sentence into an interrogative speech act.

Between the unmarked declining contour (line 3 in figure 2.10), and the level (line 1), there is room for considerable individual variation. Some speakers cluster their interrogative and nonterminal contours close to the unmarked declarative; others do the opposite and produce them just below line 1 in figure 2.10. In either case, an ordering among contours can still be observed in a trade-off with syntax: the more steeply declining marked contours typically accompany *wh*-questions, whereas the least declining contours are associated with questions with word order inversion, reminiscent of Bo's (1933) observation.

For a given intonation type, the range spanned by the contour is constant, whether it contains two or (up to) five stress groups. If it is not level, it descends gradually and evenly through stressed syllables which are equidistantly spaced in frequency, with

frequency intervals that are inversely proportional to the number of stressed syllables, as depicted in the stylized contours in figure 2.11. But their timing will depend on stress group length (in terms of the number of unstressed syllables), as hinted by the nonuniformly sloping lines between neighboring stars in the leftmost and middle contours in figure 2.11.

**2.4.10.2 Intonation cues to modality are global, not local** It is explicit in the model in figure 2.10 that no specific and elaborate $F_0$ movement is attached to the end of the contour, whether to mark a default sentence accent, a prosodic boundary, or utterance modality. Intonational markedness (less or more declarative-like) and utterance completion are inherent in the global course of the intonation contour, possibly in conjunction with the derived variation in the magnitude of stress group pattern rises, as well as diminishing intensity. This property is shared by the majority of regional Danish variants.

**2.4.10.3 Syntactic structure in simple sentences is immaterial** Within a simple sentence, the syntactic boundaries leave no trace in the course of $F_0$. The onset of each stressed vowel sets off a new $F_0$ pattern. You can move the constituents around as shown here (translated literally), and each of the resulting utterances will have basically the same course of $F_0$:

*Der går mange busser fra Tiflis* (There go many buses from Tiflis).
*Fra Tiflis går der mange busser* (From Tiflis go there many buses).
*Busser fra Tiflis går der mange af* (Buses from Tiflis go there many of).

**2.4.10.4 Longer utterances and prosodic boundaries** Some speakers manage to read a sentence with eight stressed syllables on one gradually declining intonation contour, as in the bottom panel of figure 2.5. But it is more common that above four or five stressed syllables, prosodic boundaries will be inserted, creating a succession of prosodic phrase contours within the intonation contour, as in the long utterance in figure 2.12a: *R**a**sk er s**i**kkert den **e**neste d**a**nske spr**o**gforsker hvis // n**a**vn som v**i**denskabsmand er k**e**ndt // **u**den for snævre spr**o**gvidenskabelige kr**e**dse* (Rask is probably the only Danish linguist whose name as a scientist is known outside of narrow linguistic circles). (The double slashes mark

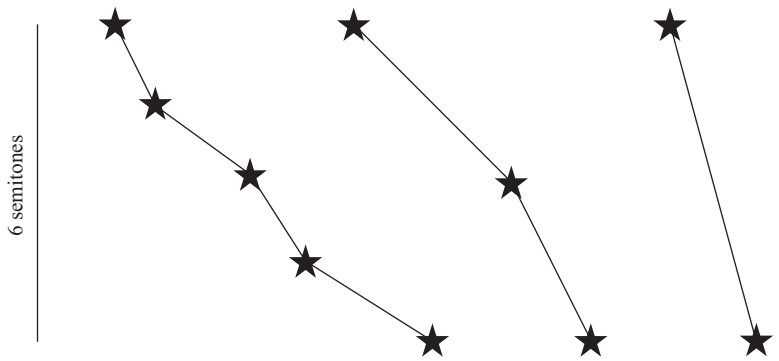


**Figure 2.11**
Stylized intonation contours with five (left), three (middle), and two (right) stressed syllables.

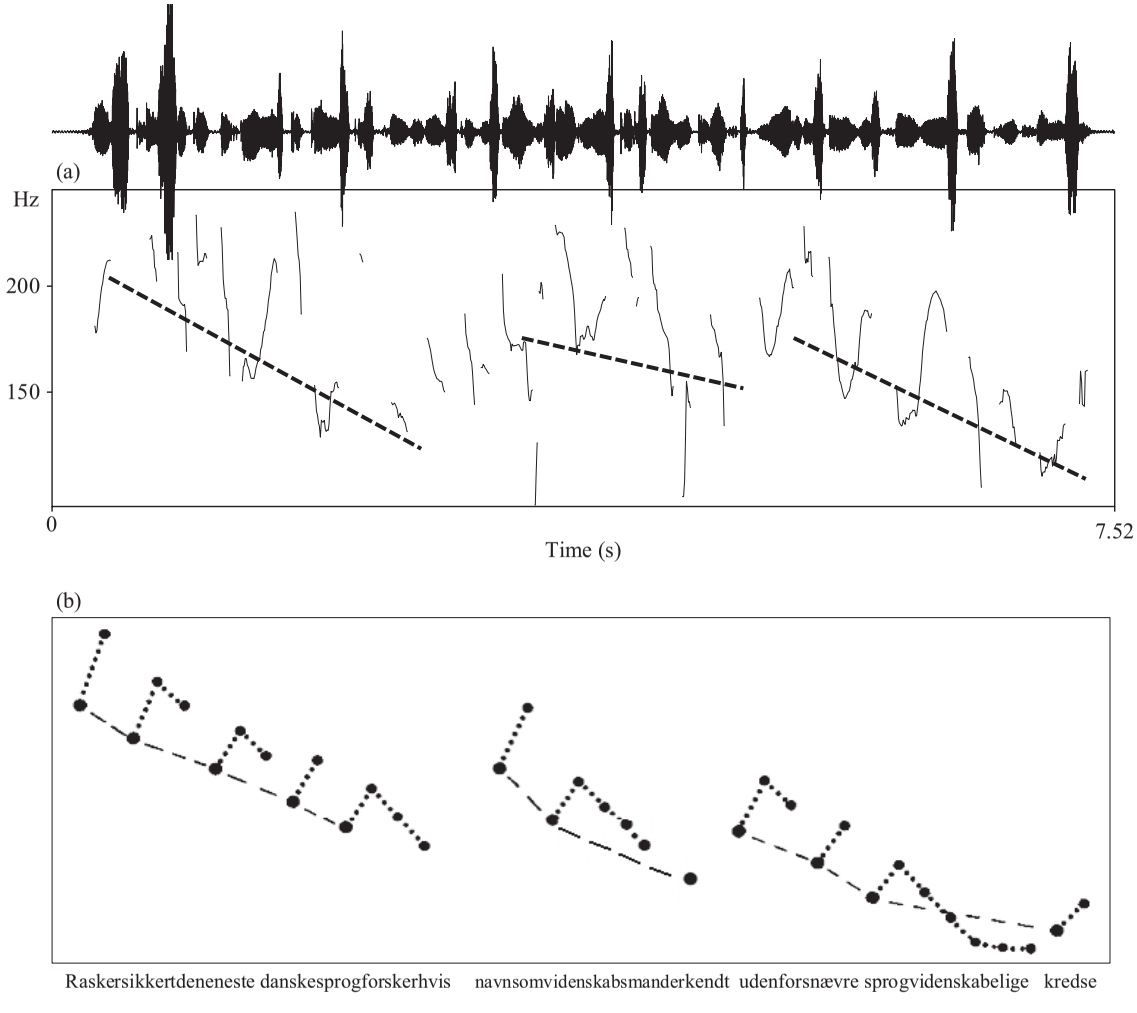


**Figure 2.12**
(a) Fundamental frequency tracing of a long utterance, *Rask er sikkert den eneste danske sprogforsker hvis // navn som videnskabsmand er kendt // uden for snævre sprogvidenskabelige kredse* (female speaker). (b) Stylized version of the fundamental frequency tracing in panel (a). Heavier dots denote stressed syllables and lighter dots unstressed syllables; broken lines denote the phrasal contours; and dotted lines the prosodic stress group patterns.

prosodic phrase boundaries.) Microprosody and perturbations due to the laryngealization in stød interfere in the $F_0$ course of the utterance, and to clarify, figure 2.12b depicts a stylized version of the utterance. This is not one smoothly declining intonation contour. Rather, the utterance has been parsed into three prosodic phrases with associated phrase contours (the broken lines), with partial resettings before the second and third phrases. Note that phrase contour onsets, defined as the frequency location of the first stressed syllable in the phrase, decline gradually through the utterance, as do offsets, defined as the frequency location of the last stressed syllable in the phrase.

One aspect of utterance-internal prosodic phrase boundaries is not in evidence in figure 2.12 (and not explicit in figure 2.13 either). The $F_0$ rise in the last stress group before an internal prosodic phrase boundary is slightly higher (by less than one semitone) than preceding $F_0$ rises in the phrase. It may be an independent gesture to signal noncompletion if it is not simply an assimilation to the higher onset of the succeeding prosodic phrase.

In the scripted material from which I derived the model of intonation presented here, the following is valid for *simple* sentences, in isolation or in combination:

1. The syntactic structure of short, simple sentences is not reflected in their intonation contour. Neither the order of constituents nor their internal structure matter.
2. Longer utterances are produced in a descending sequence of declining prosodic phrases, but the conditions governing the location of the breaks are complex: (i) A prosodic phrase must contain at least two stress groups. (ii) Prosodic phrases tend to be of equal length. But (iii) this balancing tendency is easily overruled: The prosodic boundary cannot occur within a simple first-order syntactic constituent or between semantically coherent syntactic constituents. Accordingly, *Der vil gå mange store Røde Kors busser til Grosny i Tjetjenien i aften* (Many big Red Cross buses will depart for Grozny in Chechnya this evening.) will be read (at least by fluent readers/speakers) in one sweep, in spite of its length and in spite of the formal syntactic boundary before the place complement, because buses are intimately associated semantically with their destination (or their point of departure). Thus, buses, or bus routes, may be named after their destination in one word, a compound with one main stress on the location and reduced stress on *bussen*, as in *Rombussen* (the bus to Rome).
3. Coordinate main clauses are more likely than subordinate and main clause constructions to be separated by a resetting of the intonation contour.
4. Individual intonation contours are steeper, even the text-final one, and demonstrate a greater amount of resetting between them in a succession of terminal declaratives than in a corresponding string of coordinate main clauses.

From multiple linear regression analysis of read sentences of diverse syntactic structure, consisting of up to three prosodic phrases, Petersen (1999, 2001) formulated mathematical expressions whose predictions emulated observed phrase and utterance intonation contours, confirming the global nature of utterance modality and the look-ahead involved in producing intonation contours in utterances of varying length.

### 2.4.11 An Extended Model

All of the observations noted thus far can be assembled into a more elaborate model, as in figure 2.13. The two Ts designate the overall textual contour (dotted line); U1, U2, and U3 are utterance contours (full lines), the first one of which is depicted with three potential slopes, 1, 2, and 3, corresponding to the contours in figure 2.10. P1 and P2

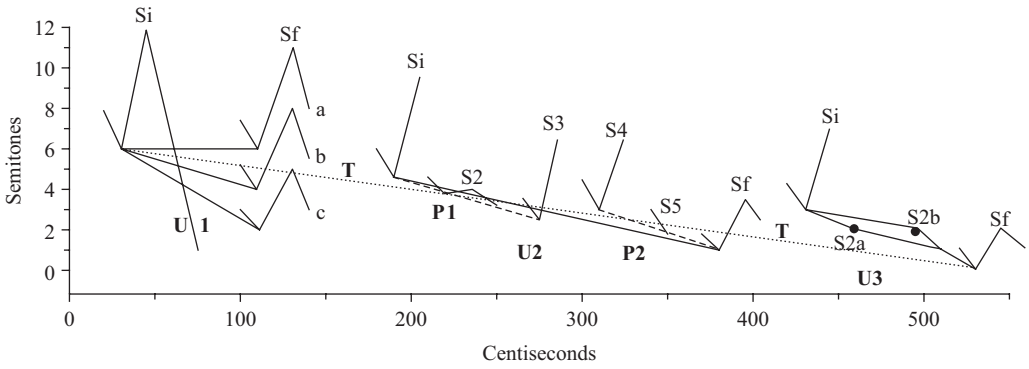


**Figure 2.13**
An extended model of Danish intonation.

are phrase contours (broken lines) within U2. The Ss designate stress group patterns (nonsaturated lines), where Si is initial in the utterance, Sf is final, and intermediate stress groups in the utterance are numbered successively from S2 and upward. The lack of $F_0$ rise in S2 (in P1/U2) is indicative of secondary stress, and S5 (in P2/U2) is a monosyllabic stress group. S2a and S2b in U3 illustrate how different duration of the initial stress group will produce different utterance contour shapes.

**2.4.11.1 The textual contour** Short texts are characterized by a gradual global declination, the textual contour, T, in figure 2.13. Its onset is defined by the first stressed syllable, Si, in the first utterance and its offset by the last stressed syllable, Sf, in the last utterance. It typically spans half an octave. Superposed on T are the individually declining utterance contours, U1, U2, and U3. The resetting of utterance contour onsets is only partial, so that utterance onsets lower gradually through the text and typically span three semitones. Utterance offsets also lower gradually but typically span only two semitones, so that utterance slopes are slightly steeper initially than finally in the text, everything else being equal (i.e., in utterances of equal length and identical modality). In texts of more than four utterances, the medial part of the textual contour will level out. It is unreasonable to expect speakers to be able to manipulate successive utterance onset descent to a degree finer than one semitone, particularly in view of their relatively large mutual temporal distance. Coordinate main clauses are less slanted relative to the global textual contour than are a sequence of terminal declaratives (not depicted in figure 2.13).

The organization of the phrase contours, P1 and P2, relative to the superordinate utterance contour, U2, mirrors the organization of utterances within the text. Phrases descend along the utterance contour. Each phrase is associated with its own slope, so the first phrase onset is higher than the last phrase onset, and the first phrase offset is higher than the last phrase offset. Intermediate phrases have intermediate onsets and offsets, but above four prosodic phrases, the utterance contour will level out medially so as not to frustrate speakers' control over step-down magnitude between phrase onsets.

**2.4.11.2 Limitations** The model's strongest limitation is its restriction to speech read aloud under laboratory conditions. Furthermore, syntactic boundaries below the sentence level are not incorporated. This is not from any a priori theoretical preclusion of their relevance, but for lack of sufficient data, and in recognition of the complexity

involved. Obviously, we need more research into the way syntactic boundaries manifest themselves in intonation.

### 2.4.12 Subordination and Superposition: Look-Ahead and Nonlocality

Why posit such a layered, hierarchical system of simultaneous, interacting, noncategorial intonational components, of varying structural and temporal scope, where larger scope components carry and set the scale for smaller scope ones? Why is implementation of $F_0$ events performed on the basis of upcoming as well as preceding events and sensitive also to syntactic and semantic structuring? Why not posit a computationally simpler abstract representation, in terms of a linear sequence of categorially different, noninteracting pitch accents whose manifestation is exclusively locally determined, implemented on a left-to-right basis? This is, in point of fact, the structure of the autosegmental-metrical theory (see Ladd 1996, 2008, and Arvaniti, chapter 1, this volume), a model in which there is no intonation component separate from those sequenced local pitch accent targets, and where global trends are only the result of iterative application of local downstep rules.

First, *pitch accent*, with its connotation of phonologically distinct differences, is not an appropriate term for stress group patterns in Danish because (i) there would be only one category, and it would always align in the same fashion with the segmental material; and (ii) its phonetic manifestation is predictable. In other words, a speaker cannot—at least not in the speech style represented by the model—make a choice between various types of pitch accent. And when, in spontaneous speech, the magnitude of stress group pattern rises is manipulated, to cue varying degrees of prominence through the utterance, we are dealing with a scalar, not a binary, phenomenon. Speakers are not making a choice in those circumstances either, of a particular pitch accent, from a set of phonologically distinct ones, but are simply subordinating the manifestation of stress to the demands for signaling more or less prominence. I have seen Bornholm speakers invert their slowly falling-rising $F_0$ patterns across a whole utterance, with a resulting change in perceived speech style or register, and I am certain a similar mechanism, a "long component" or "setting," is operative in Standard Danish. But that does not prove the existence of another pitch accent. It shows only that the manifestation of stress interacts with parameters at other levels of description.

Finally, if pitch accent were phonological, stress would not be—a proposal whose only merit is its nonconventionality. Thus, for instance, syllables with reduced stress have all the segmental characteristics of stressed syllables (full vowel quality, vowel length, segment duration, distinct articulation, stød if the phonetic and phonological conditions are satisfied) but no autonomous pitch pattern. If pitch were the distinctive factor responsible for these properties in stressed syllables, there would be no way to account for their presence in syllables with secondary stress. So again, stress group patterns are not part of the abstract representation of intonation. This is how and why, with proper scaling, the model will cover the majority of regional variants of Standard Danish as well. The principal difference among regional varieties is in the shape and the magnitude of the stress group pattern and in its alignment with segments (see figure 2.14).

The variation in shape and range of the $F_0$ patterns in figure 2.14 is rather astounding. Differences among the patterns are easier to capture when the governing principle in their execution is made explicit. In all of them, except Bornholm, there is only one point in the melodic fragment that is constrained in terms of its alignment with a segment in the prosodic stress group, as indicated by the arrows in the figure. This point is

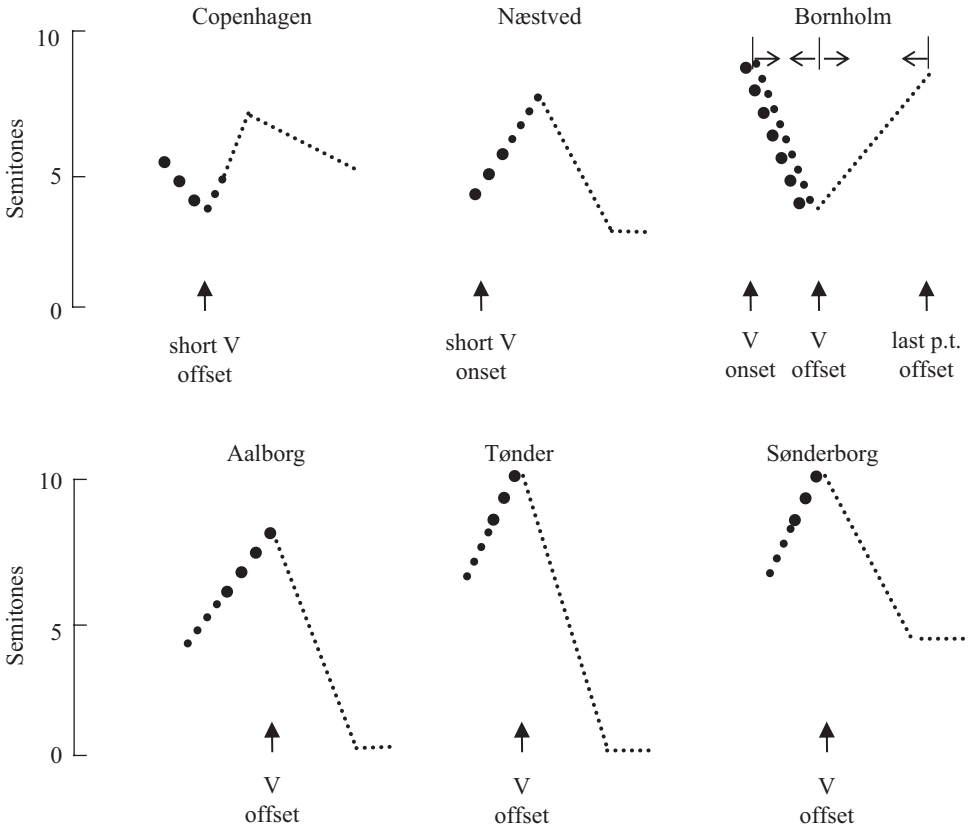


**Figure 2.14**
Model prosodic stress group patterns in six regional varieties of Standard Danish. Heavy dots depict the location of short stressed vowels on the $F_0$ pattern, medium dots depict the extension of the $F_0$ movement in long stressed vowels, and fine dots depict the course of unstressed syllables in the prosodic stress group. In Bornholm, the horizontal arrows enclose movements that may be expanded or compressed in time, in concordance with the duration of the stressed vowel (long or short–the falling part) and the number of posttonic (p.t.) syllables in the stress group (the rising part). Vertical arrows point to the segmental anchor point.

low relative to the first posttonic in Copenhagen and Næstved (on Zealand) and high in Aalborg, Tønder, and Sønderborg (in Jutland).

The anchor is invariably associated with the stressed vowel offset in Aalborg, Tønder, and Sønderborg, but with its onset in Næstved. In Copenhagen, the low turning point coincides with the offset of the stressed vowel if it is short and occurs about halfway through a long stressed vowel. Leaving out Bornholm, neither segments nor syllables in the prosodic stress group on either side of the anchor have separate tonal representations. In other words, they are simply strung out on the melody, like pearls of varying length on an undulating string. Voiceless segments interrupt the melody and will have microprosodic effects, but do not otherwise interfere with the score. A complete pattern materializes only insofar as there are syllables to carry it; otherwise it is truncated from the end. Accordingly, monosyllabic stress groups, with a short vowel

surrounded by voiceless consonants, exhibit only that part of the melody indicated by the heavy dotted line part of these five patterns in figure 2.14.

Bornholm stands out from the other varieties by the very elastic relation between segments and $F_0$. The duration of the falling part of the tonal pattern varies (albeit slightly) with the duration of the stressed vowel so that the low turning point coincides with the offset of the vowel, be it short or long. The duration of the rise varies with the duration of the unstressed part of the stress group so that the rise terminates on the last posttonic syllable. This latter variation is considerable, between one and many posttonics, and perceptually the rise is more conspicuous than the fall. In other words, the pattern has three targets: a high onset, a low turning point, and a high offset. There is, however, a limit to how fast the fall and rise may be executed, a compressibility maximum; when this limit is reached—as in a monosyllabic stress group with a short vowel surrounded by voiceless consonants—the fall disappears, and what remains is the rise from a low onset to a high offset.

Whatever the $F_0$ patterns, they are superposed on a contour defined by the stressed syllables without interfering with their frequency scaling, but the intonation contour (its location in the text, its slope, its length), on the contrary, is decisive for the manifestation of stress group patterns. Furthermore, the manifestation of prominence degrees above normal stress demonstrates yet more syntagmatic interaction. Stress group patterns are reduced before emphasis for contrast, but not before a mere focus, that is, the realization depends on the nature of the succeeding stress group.

Furthermore, utterance and text intonation contours cannot be computed on a purely left-to-right basis either. Slope varies in inverse proportion to length. Stressed syllables in an utterance are more closely spaced in frequency in longer than in shorter utterances. In other words, the location of a stressed syllable in the frequency range is sensitive to the number of succeeding ones. For example, the second stressed syllable is higher in an utterance of five than one of three stressed syllables, everything else being equal, as depicted in figure 2.11. A similar look-ahead is operative in the combination of utterances into prosodically coherent texts. In this manner, utterance and text contours are subject to compression and expansion in time, a process that is inconceivable without look-ahead. Pre-planning also turns up in the difference between coordinate main clauses, which are less slanted relative to the global textual contour than a sequence of terminal declaratives. Look-ahead and interaction turn up in the temporal structure as well: speaking rate is somewhat accelerated before a (noninitial) focused item.

Finally, intonation contours, except the maximally marked ones, are always associated with some degree of declination. Thus, for instance, there are no conventional declaratives without declination.

These facts should make an account of sentence intonation in Danish, in terms of (varying degrees of) local downstep or range reduction, triggered by certain pitch accent configurations, perhaps a formal possibility but empty of the significance it carries in tone languages. Note that the global character and the hierarchical principle underlying the superposition and subordination do not exclude the existence of features of a singular and noninteracting character such as boundary signals. Thus, there are Danish regional varieties that definitely have a specific, final boundary tone that does not interact with what precedes it, as in Sønderborg and Bornholm. Likewise, a hierarchical organization does not preclude lexical tonal differences.

To sum up, the hierarchical and sequential models do not differ in the acknowledgment of look-ahead, but they differ in its representation. For example, in linear

sequence models, longer utterance onsets are higher than shorter ones, but pitch accent relations are unaffected. In the hierarchical approach, utterance onsets need not vary, but the declining slope is less steep in longer than in shorter utterances. Liberman and Pierrehumbert (1984) stated that a pitch accent can only look back to a previous pitch accent, a phrase to a previous phrase, and so on, and apparent instances of anticipation should be explained by, for example, feature spreading or temporal overlap in the realization of the segments in question (which is really only passing the buck). The hierarchical representation entails a more direct interaction between subconstituents and superordinate structures.

### 2.4.13 Intonation and Meaning

The model described here has only two components: the prosodic stress group pattern and the superordinate intonation contour, with its potential subordinate phrase contours. The prosodic stress group pattern lends prominence to stressed syllables, but has no meaning in and of itself. The only potential candidate, then, for a direct link to meaning, albeit of a nonlexical nature, is the intonation contour. And indeed, figure 2.10 pointed to a correlation between intonation contour slope and sentence type. For instance, a declarative sentence in conventional use is typically associated with the most declining intonation contour: one that falls to the bottom of the speaker's normal speaking range. At the other extreme, declarative sentences that function as interrogatives are typically associated with a nondeclining intonation contour. But that is not tantamount to saying that any of these contours, in and of themselves, carry meaning in any narrow sense, at least not in linguistically coded messages. No one intonation contour invariably "means" one thing only, or invokes just one modality, just as no one modality is invariably associated with any one intonation contour. The larger context determines how any given sentence, with any given intonation contour, is to be interpreted, taking also unconventional and nontypical usage into account.

Some examples follow in table 2.1 from a corner of a phonetician's universe, all about what phoneticians do and what it takes to be one. They involve questions versus statements, and they demonstrate that any sentence, with any intonation contour, may be interpreted as either a question or a statement, according to the wider context. The numbers in parentheses in the intonation contour column refer to the contour slopes in figure 2.10.

Here are staged but plausible wider contexts for two of the examples in table 2.1. (i) The declarative sentence that opens an interview: A famous phonetician, now deceased, is invited to be interviewed on the Danish radio. The journalist opens the interview with a statement, with a typical unmarked statement intonation contour: *Du kalder dig selv fonetiker* (You call yourself a phonetician). But the utterance is correctly interpreted by the interviewee as a request for confirmation and an invitation for her to talk about herself and her life as a phonetician. (ii) The interrogative sentence that is interpreted as a polemical statement: A historian of science is being interviewed about various prominent figures in the humanities in Denmark and calls one of them a linguist, but then goes on to mention her important analysis of the acoustics of stop consonants. The journalist responds with a question: *Hvad er en fonetiker?* (What is a phonetician?) with a typical question intonation contour, but is being polemical and actually stating that with that kind of activity, one is a phonetician, not a linguist. Readers are invited to use their imagination to create contexts for the remaining six examples.

**Table 2.1**
Examples of declarative and interrogative sentences with different intonation contours and their interpretation in isolation and in various staged contexts

| | The spoken utterance | | |
|---|---|---|---|
| The sentence | Intonation contour | Interpretation in isolation | Interpretation in various contexts |
| Declarative: *Du kalder dig selv fonetiker* (You call yourself a phonetician.) | Nondeclining (1) | Question | *Question*: Why not a linguist? |
| | | | *Ironical statement*: You are NOT a phonetician. |
| | Declining (3) | Statement | *Question*: Opening an interview |
| | | | *Statement*: And yet someone else might call you a LINGUIST. |
| Interrogative *Hvad er en fonetiker* (What is a phonetician?) | Nondeclining (2) | Question | *Question*: True request for information. |
| | | | *Polemical statement*: That makes her a PHONETICIAN. |
| | Declining (3) | Statement | *Polemical question*: If SHE is not a phonetician, then what IS a phonetician? |
| | | | *Statement in conclusion*: That makes her a phonetician. |

### 2.4.14 Transcription

The model was developed to account for speakers' production of short utterances and texts, in a distinct and formal style of speech, not with any specific application to transcription in mind, whether in fieldwork or didactic application. But its basic tenets—the stress group pattern and the intonation contour—have served well as anchor points in the transcription of spontaneous Danish speech in the DanPASS corpus (Grønnum 2009).

The corpus is segmented down to word and syllable levels wherever possible. Orthography in tier 1 in the Praat screen shot in figure 2.15 (Boersma and Weenink 2006) is supplemented with information about stress, in the shape of commas before the vowels of the stressed syllables; word class information is fully specified in tier 2 and simplified in tier 3; the phoneme representation is in tier 4; an idealized pronunciation in tier 5; the actual phonetic transcription, relatively fine grained, at the word level in tier 6; and broken into syllables in tier 7. Tiers 8 and 9 hold the prosodic information. Stressed syllables are marked with a star. If more than normally/neutrally prominent, the star is preceded by an exclamation mark. The magnitude of the rise to the first posttonic syllable may be higher than normal (H/; not evidenced in figure 2.15), simply high (H), or smaller than normal (h). If a posttonic syllable, for whatever reason, is lower than the preceding stressed syllable, it is marked in analogous fashion with L\ (much lower), L (lower), and l (a little lower), respectively (not evidenced in figure 2.15 and not very common). In tier 9, the perceived pitch level of each stressed syllable in

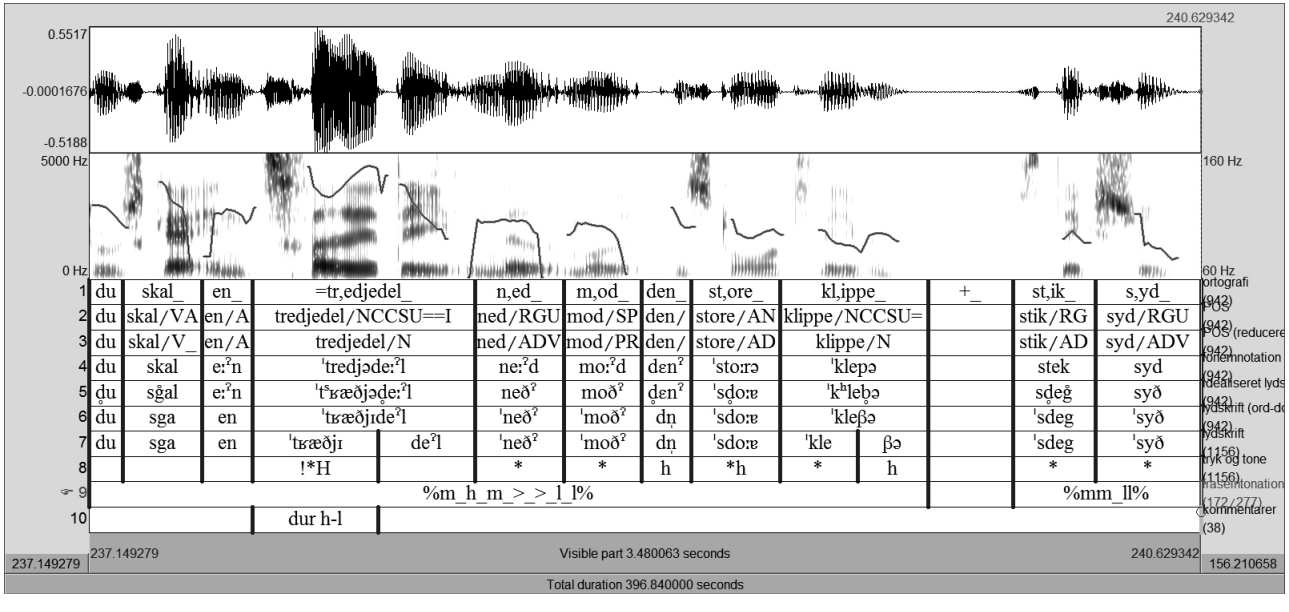


**Figure 2.15**
Praat screen shot from the DanPASS corpus. *Source*: http://www.danpass.hum.ku.dk.

the phrase is annotated to three degrees: high (h), mid (m), and low (l), respectively. Gradual decline over several stressed syllables, for instance, from mid to low, is marked with arrows, as shown in the figure. No attempt has been made to relate contours across phrase boundaries, but the information may easily be supplied if required.

Note that any similarity with the tones and break indices (ToBI) convention for transcribing prosody (see Jun, chapter 4, this volume) is merely superficial. For the description of Danish intonation, the phonological assumptions behind ToBI are inappropriate, and as a phonetic transcription system, it is not sufficiently fine grained. (For a more detailed discussion, see Thorsen 1985, 1986a, 1986b, and Grønnum 1995.)

### 2.4.15 Psychological Reality

A linear representation is computationally somewhat simpler, and thus easier to implement in rule synthesis, than layered structures involving look-ahead, but that is not necessarily indicative of how speakers and listeners process intonation. And look-ahead is definitely operative and evidenced in other aspects of speech, so why not in intonation?

We are all both speakers and hearers, and although we do not a fortiori produce and perceive in the same terms, it is at least not an unlikely assumption. I hypothesize that utterance intonation in Danish is holistically conceived, and that this concept is cognitively simpler to the speaker and listener than the summation of its atomic elements, that is, the local ups and downs. Anecdotal evidence hints that this hypothesis would be worth testing. When linguistically naïve Danes are asked to characterize the melody of various Danish utterances, they typically provide gross overall shapes. They have to be pushed hard—and you have to exaggerate the patterns—for them to hear that there are local pitch movements associated with stressed syllables. The local humps are not conceived as part of the melody, and listeners seem to disregard the contribution to pitch that comes from stress.

## 2.5 From Read to Spontaneous Speech

What happens when the model is confronted with spontaneous speech? I do not expect look-ahead and preplanning to be much in evidence. If the speaker does not know in advance what he is going to say, but proceeds as his thoughts develop, he may not only have to insert prosodic phrase boundaries and pauses in suitable places, but he may also be forced to lower the utterance offset in longer utterances relative to shorter ones. Nor do I expect to encounter the even distribution of prominence that characterizes isolated read utterances. On the contrary, prominence will vary among the stressed words as the pragmatic context dictates. Therefore, $F_0$ rises in stress group patterns will vary in magnitude, with higher rises after more prominent syllables (see section 2.4.5; Jensen and Tøndering 2005a, 2005b), and global contours will not progress in smooth and regular fashion because prominence beyond neutral may propel the stressed syllable upward in the frequency range, making the contour bumpy. On the other hand, I would still expect utterance modality to be inherent in the global trend of its contour and stress group $F_0$ patterns to be superposed on the contour without interfering with its execution, except for the timing inherent in the duration of individual stress groups.

A handful of studies of spontaneous Danish address these issues. Dyhr (1993) analyzed a total of 302 utterances from three speakers, interviewed in their homes. He mostly confirms the validity of the model's assumptions but also finds that longer declarative utterance contours may level out toward the end or perform a rising-falling movement from start to finish. Dyhr (1995) notes that emphasis may be achieved not

only by lifting the stressed syllable upward in the frequency range, but occasionally also by lowering it, relative to surrounding stressed syllables.

Grønnum and Tøndering (2007) analyzed three hundred questions of various types, plus 51 declarative utterances for comparison, from some of the dialogues in the DanPASS corpus. In spite of considerable variability in the material, there is nevertheless a tendency for different global trends among the contours, *and* an association between global trend and utterance modality, although with a slightly different ordering than in read speech. Questions with word order inversion are slightly less declining than declarative questions. Surprisingly, final $F_0$ rises were somewhat higher in the declarative utterances and smaller on the less declining interrogative contours, contrary to read speech. However, new information is typically found last in declaratives but first in questions (Lambrecht 1994), which may account for the distribution of $F_0$ rises in these dialogues.

Tøndering (2008, 2010) provides a thorough analysis of intonation and syntax in nearly all the monologues in DanPASS: 2,604 prosodic phrases, from fifteen speakers, containing between zero and fifteen stressed syllables. He encountered 550 different sequences of the four symbols in the annotation (*h*, *m*, *l*, *>*). For example, there are 469 prosodic phrases with four stressed words; they occur with 91 different contours (of 111 theoretically possible ones). No model can be expected to predict this kind of variability, and 'a model of intonation in spontaneous speech' is actually a contradiction in terms. But we may inquire which features of read speech recur in spontaneous speech, though we may not be able to quantify them accurately. Tøndering finds no evidence of look-ahead and preplanning in the data, in terms of equidistantly frequency scaled stressed syllables in contours of constant range. On the contrary, longer phrases have lower offsets than shorter ones do, as expected. Nor is the magnitude of the $F_0$ rise in stress group patterns found to decrease gradually through the phrase, also as expected. On the other hand, Tøndering reports no evidence that $F_0$ patterns per se—by virtue of their peak or offset frequency—have any role in the scaling of the stressed syllables. And in spite of the rather forbidding variability, he nonetheless discovers an overall declining trend in the bumpy intonation contours. (There are hardly any questions in these monologues.) He finds, furthermore, that $F_0$ rises are highest in the last stress group in the phrase. That is where new information is located, hence with more prominence, hence higher $F_0$ rises.

In summary, applied to the description of intonation in spontaneous speech, with its contextually and pragmatically induced variation in prominence, the model needs to adjust the smooth progression of the stressed syllables in prosodic phrases and utterances, as well as the gradual reduction of the associated $F_0$ patterns. Insofar as final stressed syllables are the more prominent ones in declaratives, that will, in and of itself, make final $F_0$ rises higher than preceding ones and also higher than in nondeclaratives, where new information tends to occur early, not late as in declaratives.

## 2.6 Conclusion

If any feature in Danish prosody can lay claim to fame for its complexity, it would be the stød (Basbøll 2005; Grønnum, Vazquez-Larruscaín, and Basbøll 2013), not intonation, which could hardly be simpler. What most obviously distinguishes standard Danish intonation from other languages—notably its nearest neighbor, Swedish—is the small inventory of components. There are only two: a global utterance contour, which may be partitioned into a succession of subordinate prosodic phrase contours and may

also combine to form overarching textual contours, and a local $F_0$ pattern associated with the prosodic stress group. In spite of its considerable variability, intonation in spontaneous speech can be accounted for with the same two entities. The difference in final $F_0$ rise magnitude between declaratives and interrogatives could have been a candidate for a separate cue to modality if it were not for the contraintuitive fact that these rises are more extensive in declaratives than in interrogatives. This difference has to do not with modality per se but with the different way information is structured in questions and statements.

The other distinguishing feature is what in read speech appears as a hierarchical organization of the two components: $F_0$ patterns are superposed on and subordinate to the intonation contour. Superposition means that stress group patterns do not interfere with the intonation contour as defined by the stressed syllables. Subordination means that the magnitude of $F_0$ excursion in the stress group pattern depends partly on the kind of contour on which it rides (higher rises on less declining contours) and partly on its location on the contour (rises shrink from start to finish). The need to highlight some words in spontaneous speech and relegate others to the background introduces a gradation of acoustic and perceptual prominence that wreaks havoc with the smooth intonation contours of read speech. Furthermore, with more or less prominence, $F_0$ patterns are more or less extensive. Accordingly, they do not shrink gradually toward the end, and their subordination to the utterance contour is not discernible. But as long as the frequency scaling of the stressed syllables is unaffected by the associated $F_0$ patterns in terms of their peak or offset frequency, the intonation contour is still superordinate to the manifestation of the prosodic stress group pattern, and a hierarchy can be said to exist.

Evidently, the last word has not been said about Danish intonation, and many more analyses of spontaneous speech are awaiting future investigations.

## Acknowledgments

Many thanks to Hans Basbøll for extensive and insightful comments on a draft of this chapter.

## Note

**1**. Until the mid-1980s I published under my married name as Nina Thorsen and until 1989 as Nina Grønnum Thorsen.

## References

Andersen, P. 1954. "Dansk Fonetik." In *Nordisk Lærebog for Talepædagoger*, 308–354. Copenhagen: Rosenkilde og Baggers Forlag.

Arnholtz, A. 1939–1940. "Talekunstens og talelærens udvikling og stilling i Danmark." *Nordisk Tidsskrift for Tale og Stemme* 4:82–98.

Arnholtz, A. 1942. "Om vulgær-københavnskens melodik." *Nordisk Tidsskrift for Tale og Stemme* 6:23–24.

Arnholtz, A. 1954. "Dansk Foredragslære." In *Nordisk Lærebog for Talepædagoger*, 422–451. Copenhagen: Rosenkilde og Baggers Forlag.

Basbøll, H. 2005. *The Phonology of Danish*. Oxford: Oxford University Press.

Bloch, S. N. J. 1805. *Forsøg til en Theorie for det udvortes Foredrag*. Odense, Denmark: S. Hempel.

Bo, A. 1933. *Tonegangen i Dansk Rigsmål*. Copenhagen: Branner.

Boersma, P., and D. Weenink. 2006. *Praat: Doing Phonetics by Computer*. Computer program praat5404 (accessed December 29, 2014, at http://www.praat.org/).

Brink, L., and J. Lund. 1975. *Dansk Rigsmål 1–2, Lydudviklingen siden 1840 med særligt henblik på sociolekterne i København*. Copenhagen: Gyldendal.

Bruce, G. 1977. *Swedish Word Accents in Sentence Perspective*. Lund, Sweden: Gleerup.

Dyhr, N. J. 1993. "An Acoustical Investigation of the Fundamental Frequency in Danish Spontaneous Speech." In *Nordic Prosody VI*, edited by B. Granström and L. Nord, 23–32. Stockholm: Almqvist and Wiksell.

Dyhr, N. J. 1995. "The Fundamental Frequency in Danish Spontaneous Speech with Special Reference to Syllables Boosted for Emphasis." In *Aspects of Danish Prosody*, edited by J. Rischel and H. Basbøll, 49–67. Odense, Denmark: Odense University Press.

Fischer-Jørgensen, E. 1946, 1947. Unpublished lecture notes (now belonging to the author).

Fischer-Jørgensen, E. 1983. "The Acoustic Manifestation of Stress in Danish with Particular Reference to the Reduction of Stress in Compounds." In *Sound Structures: Studies for Antonie Cohen*, edited by M. van den Broecke, V. van Heuven, and W. Zonneveld, 81–104. Dordrecht, the Netherlands: Foris Publications. A more comprehensive version appears in *Annual Report of the Institute of Phonetics, University of Copenhagen* 18 (1984): 45–161.

Forchhammer, G. 1903. *Om Nødvendigheden af Sikre Meddelelsesmidler i Døvstummeundervisningen*. Copenhagen: J. Frimodts Forlag.

Fujisaki, H., K. Hirose, and K. Ohta. 1979. "Acoustic Features of the Fundamental Frequency Contours of Declarative Sentences in Japanese." *Annual Bulletin—Research Institute of Logopedics and Phoniatrics* 13:163–173.

Gårding, E. 1979. "Sentence Intonation in Swedish." *Phonetica* 36:207–215.

Grønnum, N. 1990. "Prosodic Parameters in a Variety of Regional Danish Standard Languages, with a View towards Swedish and German." *Phonetica* 47:182–214.

Grønnum, N. 1991. "Terminality and Completion in Danish, Swedish and German." In *Proceedings of the Twelfth International Congress of Phonetic Sciences*, edited by the Congress Committee and the Organizing Committee, vol. 4, 270–273. Aix-en-Provence, France: Université de Provence, Service des Publications.

Grønnum, N. 1992. *The Groundworks of Danish Intonation: An Introduction*. Copenhagen: Museum Tusculanum Press.

Grønnum, N. 1993. "Perceptual Invariance in Danish Stress Group Patterns." In *Nordic Prosody VI*, edited by B. Granström and L. Nord, 77–85. Stockholm: Almqvist and Wiksell.

Grønnum, N. 1994. "Rhythm, Duration and Pitch in Regional Variants of Standard Danish." *Acta Linquistica Hafniensia* 27:190–218.

Grønnum, N. 1995. "Superposition and Subordination in Intonation: A Non-Linear Approach." In *Proceedings of the Thirteenth International Congress of Phonetic Sciences*, edited

by K. Elenius and P. Branderud, vol. 2, 124–131. Stockholm: Stockholm Congress Organisers at Kungliga Tekniska Högskolan and Stockholm University.

Grønnum, N. 1998a. "A Critical Remark to D. R. Ladd's Intonational Phonology." *Journal of Phonetics* 26:109–112.

Grønnum, N. 1998b. "Intonation in Danish." In *Intonation Systems: A Survey of Twenty Languages*, edited by D. Hirst and A. di Cristo, 131–151. Cambridge: Cambridge University Press.

Grønnum, N. 2000. "Danish: Illustrations of the IPA." *Journal of the International Phonetic Association* 28:99–105.

Grønnum, N. 2007. *Rødgrød med Fløde. En lille bog om dansk fonetik*. Copenhagen: Akademisk Forlag.

Grønnum, N. 2009. "A Danish Phonetically Annotated Spontaneous Speech Corpus (DanPASS)." *Speech Communication* 51:594–603. https://doi:10.1016/j.specom.2008.11.002. The corpus is available at www.danpass.hum.ku.dk.

Grønnum, N., and J. Tøndering. 2007. "Question Intonation in Non-Scripted Danish Dialogues." In *Proceedings of the Sixteenth International Congress of Phonetic Sciences*, edited by J. Trouvain and W. J. Barry, 1229–1232. Saarbrücken, Germany: Saarland University.

Grønnum, N., M. Vazquez-Larruscaín, and H. Basbøll. 2013. "Danish Stød: Laryngealization or Tone." *Phonetica* 70:66–92.

Heger, S. 1981. *Sprog og Lyd: Elementær Dansk Fonetik*. Copenhagen: Akademisk Forlag.

Hjelmslev, L. 1951. "Grundtræk af det danske udtrykssystem med særligt henblik på stødet." *Selskab for Nordisk Filologi. Årsberetning for 1947-49-50*:12–24. English translation by Francis Whitfield, "Outline of the Danish Expression System with Particular Reference to the Stød." In L. Hjelmslev, *Essais linguistiques II* (= *Travaux du Cercle Linguistique de Copenhague*, vol. 14), edited by N. Ege, E. Fischer-Jørgensen, K. Togeby, and F. J. Whitfield, 1979, 247–266. Copenhagen: Nordisk Sprog og Kulturforlag.

Hjelmslev, L. 1954. "Almindelig Fonetik." In *Nordisk Lærebog for Talepædagoger*, 233–307. Copenhagen: Rosenkilde og Baggers Forlag.

Høysgaard, J. P. 1747. *Accentuered og Raisonnered Grammatica*. In *Danske Grammatikere, Bind IV*, edited by H. Bertelsen, 1920, 249–487. Copenhagen: Nordisk Forlag. Reprint 1979, Copenhagen: Det Danske Sprog og Litteratur Selskab.

Høysgaard, J. P. 1769. *Første Anhang til den Accentuerede Grammatika*. In *Danske Grammatikere, Bind V*, edited by H. Bertelsen, 1923, 507–550. Copenhagen: Nordisk Forlag. Reprint 1979, Copenhagen: Det Danske Sprog og Litteratur Selskab,

Jensen, C., and J. Tøndering. 2005a. "Choosing a Scale for Measuring Perceived Prominence." In *Proceedings of 6th Interspeech 2005 and 9th European Conference on Speech Communication and Technology (EUROSPEECH)*, 2385–2388. International Speech Communication Association (ISCA).

Jensen, C., and J. Tøndering. 2005b. "Perceived Prominence and Scale Types." In *Proceedings of the XVIIIth Swedish Phonetics Conference*, edited by A. Eriksson and J. Lindh, 111–114. Göteborg, Sweden: Göteborg University.

Jensen, J. 1833. *Forsøg til en Dansk Sproglære*. Copenhagen: Den Gyldendalske Boghandel.

Jerndorff, P. 1897. *Om Oplæsning*. Copenhagen: Nordisk Forlag.

Jerndorff, P. 1911. *Tale og Læsning*. Copenhagen: Nordisk Forlag.

Jespersen, O. 1897–1899. *Fonetik*. Copenhagen: Det Schubotheske Forlag.

Jespersen, O. 1934. *Modersmålets Fonetik*. Copenhagen: Gyldendal.

Ladd, D. R. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladd, D. R. 2008. *Intonational Phonology*. 2nd ed. Cambridge: Cambridge University Press.

Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.

Liberman, M. Y., and J. Pierrehumbert. 1984. "Intonational Invariance under Changes in Pitch Range and Length." In *Language Sound Structure*, edited by M. Aronoff and R. T. Oehrle, 157–233. Cambridge, MA: MIT Press.

Lützen, Peter H. 2010. *Det Sproglige i Dansk*. Copenhagen: Dansklærerforeningens Forlag.

Mikkelsen, K. 1894. *Dansk Sproglære*. Copenhagen: Lehmann and Stages Forlag.

Möbius, B. 1993. *Ein quantitatives Modell der deutschen Intonation*. Tübingen, Germany: Max Niemeyer Verlag.

Öhman, S. E. G. 1968. "A Model of Word and Sentence Intonation." *Speech Transmission Laboratory: Quarterly Progress and Status Report*: 6–11. Stockholm: Royal Institute of Technology.

Oxenvad, E. 1937. *Vort Sprog, Bind I*. Copenhagen: Nordisk Forlag.

Petersen, N. R. 1999. "Modelling Danish Sentence and Phrase Intonation." In *Proceedings of the Fourteenth International Congress of Phonetic Sciences*, edited by J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey, 925–928. Berkeley: University of California Press.

Petersen, N. R. 2001. "Modelling Fundamental Frequency in First Posttonic Syllables in Danish Sentences." In *Proceedings of Eurospeech 2001 Scandinavia*, edited by P. Dalsgaard, B. Lindberg, and H. Benner, 939–943. Aalborg, Denmark: Center for Personkommunikation, Aalborg University.

Rischel, J. 1980. "Phrasenakzent als Signal des Objekts ohne 'determiner' im Dänischen." *Kopenhagener Beiträge zur germanistischen Linguistik, Sonderband 1*: 262–279.

Rischel, J. 1983. "On Unit Accentuation in Danish—and the Distinction between Deep and Surface Phonology." Special issue on prosody, edited by H. Basbøll. *Folia Linguistica* 17:51–197.

Sneedorff-Birch, F. 1835 *Dansk Retskrivningslære*. Copenhagen: J. H. Schubothe.

't Hart, J., and R. Collier 1979. "On the Interaction of Accentuation and Intonation in Dutch." In *Proceedings of the Ninth International Congress of Phonetic Sciences*, edited by N. Grønnum, E. Fischer-Jørgensen, and J. Rischel, vol. 2, 395–402.

Thorsen, N. 1978. "An Acoustical Investigation of Danish Intonation." *Journal of Phonetics* 6:151–176.

Thorsen, N. 1979. "Interpreting Raw Fundamental Frequency Tracings of Danish." *Phonetica* 36:57–78.

Thorsen, N. 1980a. "Intonation Contours and Stress Group Patterns in Declarative Sentences of Varying Length in ASC Danish." *Annual Report of the Institute of Phonetics, University of Copenhagen* 14:1–29.

Thorsen, N. 1980b. "Neutral Stress, Emphatic Stress, and Sentence Intonation in Advanced Standard Copenhagen Danish." *Annual Report of the Institute of Phonetics, University of Copenhagen* 14:121–205.

Thorsen, N. 1980c. "A Study of the Perception of Sentence Intonation: Evidence from Danish." *Journal of the Acoustical Society of America* 67:1014–1030.

Thorsen, N. 1980d. "Word Boundaries and $F_0$ Patterns in Advanced Standard Copenhagen Danish." *Phonetica* 37:121–133.

Thorsen, N. 1981. "Intonation Contours and Stress Group Patterns in Declarative Sentences of Varying Length in ASC Danish: Supplementary Data." *Annual Report of the Institute of Phonetics, University of Copenhagen* 15:13–47.

Thorsen, N. 1982a. "On the Variability in $F_0$ Patterning and the Function of $F_0$ Timing in Languages Where Pitch Cues Stress." *Phonetica* 39:302–316.

Thorsen, N. 1982b. "Selected Problems in the Tonal Manifestation of Words Containing Assimilated or Elided Schwa." *Annual Report of the Institute of Phonetics, University of Copenhagen* 16:37–100.

Thorsen, N. 1983a. "Standard Danish Sentence Intonation: Phonetic Data and Their Representation." Special issue on prosody, edited by H. Basbøll. *Folia Linguistica* 17:187–220.

Thorsen, N. 1983b. "Two Issues in the Prosody of Standard Danish." In *Prosody: Models and Measurements*, edited by A. Cutler and D. R. Ladd, 27–38. Berlin: Springer-Verlag.

Thorsen, N. 1984a. "$F_0$ Timing in Danish Word Perception." *Phonetica* 41:17–30.

Thorsen, N. 1984b. "Variability and Invariance in Danish Stress Group Patterns." *Phonetica* 41:88–102.

Thorsen, N. G. 1985. "Intonation and Text in Standard Danish." *Journal of the Acoustical Society of America* 77:1205–1216.

Thorsen, N. G. 1986a. "Intonation and Text in Standard Danish: With Special Reference to the Abstract Representation of Intonation." In *Proceedings of the Fifth International Phonology Meeting*, edited by S. L. Fulmer, W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer, and J. R. Rennison, 301–309. Cambridge: Cambridge University Press.

Thorsen, N. G. 1986b. "Sentence Intonation in Textual Context: Supplementary Data." *Journal of the Acoustical Society of America* 80:1040–1047.

Thorsen, N. G. 1987. "Suprasegmental Transcription." In *Probleme der phonetischen Transkription*, edited by A. Almeida and A. Braun, 79–109. Stuttgart, Germany: Franz Steiner Verlag.

Thorsen, N. G. 1988. "Default Sentence Accents and Focal Sentence Accents." In *Papers from the Second Swedish Phonetics Conference, May 1988*, edited by S. A. J. Wood, 120–123. Lund, Sweden: Lund University, Department of Linguistics.

Tøndering, J. 2008. *Skitser af prosodi i spontant dansk*. Copenhagen: Det Humanistiske Fakultet, Copenhagen University.

Tøndering, J. 2010. "Prosodiske fraser og syntaktisk struktur i spontan tale." *Nydanske Studier* 39:166–198.

Vaissière, J. 1983. "Language-Independent Prosodic Features." In *Prosody: Models and Measurements*, edited by A. Cutler and D. R. Ladd, 53–66. Berlin: Springer-Verlag.