From Communicative Functions to Prosodic Forms

Klaus J. Kohler
University of Kiel
Kiel, Germany

Short title: From communicative functions to prosodic forms

Prof. Dr. Klaus J. Kohler
Holm 4
D-24113 Molfsee
+49 431 651453
kkj AT ipds DOT uni-kiel DOT de
Abstract
This is a proposal in favour of proceeding from communicative function to linguistic form, rather than the reverse, for an insightful account of how humans communicate by speech in languages. A functional framework is developed that encompasses argumentation structures, declarative and interrogative functions, and expressive intensification. Such a function orientation can become a powerful tool in comparative prosodic research across the world’s languages. The potential of this approach is shown by comparing the prosodic form of Mandarin Chinese data collected in functionally contextualized scenarios with corresponding data from English and German.
1 Analysing Prosody in Speech Communication

1.1 Prosody in Structural Linguistics

The core of today’s mainstream structural linguistics is linguistic form, abstracted from phonetic substance, and represented by discrete systemic units and their structural relations in syntax, morphology and phonology for the differentiation of meaning in the grammars of languages. In this paradigm, discrete formal contrasts constitute the pivot to which both phonetic manifestation and semantic interpretation, generally restricted to propositional meaning, are referred post hoc. Therefore graded differences between speech elements, as well as attitudinal and expressive meaning, are relegated to paralinguistics, i.e. outside linguistics proper, and communicative functions in speech interaction are conceptualized as being subsidiary to linguistic form. This also excludes a direct link between communicative function and the physics of speech behaviour. These restrictions of formal linguistics have had far-reaching effects on the shaping of research into prosody. In the course of the past three decades, prosody has, on the one hand, been incorporated in formal phonology, resulting in current mainstream Autosegmental Metrical (AM) Phonology. On the other hand, prosodic analysis has come to be based on experimental data collection and primarily $F_0$ measurement in a Laboratory Phonology framework, with little or no concern for auditory evaluation by the investigator.

This formal linguistic perspective of prosody has thus moved away from the long tradition of English intonation analysis in the London School of Phonetics [Allen 1954; Armstrong and Ward 1931; Jones 1956; Kingdon 1958; O’Connor and Arnold 1961; Palmer 1924; Palmer and Blandford 1939], which relied on auditory observation and introspection for the practical application in teaching English as a foreign language, where the functions of pitch patterns played a central role. The form-centred analysis is also diametrically opposed to Bolinger [1986], who provides a wealth of fine auditory prosodic detail with great insight into communicative functions. His functional gestural approach is captured by the following quotations:

“… the approach followed here … differs from others mainly in its insistence on the independence of intonation from grammar … Intonation has more in common with gesture than with grammar …, though both gesture and intonation are tremendously important to grammar, as their lines intersect.” (p. viii)

“Nevertheless, though intonation is indispensible to grammar, the grammatical functions of intonation are secondary to the emotional ones; speakers feel differently about what they say, and
the feelings manifest themselves in pitch changes that serve as cues. One proof of the emotional rather than logical nature of intonation – its symptomatic more than symbolic character – is the fact that speakers rarely if ever objectify the choice of an intonation pattern; they do not stop and ask themselves “Which form would be best here for my purpose?” as they frequently do in selecting a word or a grammatical construction. Instead, they identify the feeling they wish to convey, and the intonation is triggered by it.” (p. 27)

“The affective, attitudinal, emotive side of intonation is inextricably intertwined with the grammatical side. There is no way to set aside what is “ideophonic” just to make the grammarian’s task easier.” (p. 34)

“The dynamic nature of up-down motion suggests a system in which processes bind together in their bearing on goals: the same goal can be achieved by different parts of the system, different neuromuscular assemblies [jaw, corners of the mouth, eyebrows, hands, shoulders]… We can … say that “speech and limb” is an artificial separation to begin with and that the overintellectualization of speech by casting all its manifestations in the mold of syntactic and morphological abstraction has obscured the true nature of intonation.” (pp. 201f)

“[He’s eating it?, uttered in ‘delighted surprise’, is pronounced] with the desired [rising] intonation and with the unconsciously adopted smile and perhaps also a forward thrust of the head suitable for the simulated question, The effect on the resonators is audible. The subjects then react to the whole gestural complex, not just to the intonation. Visible gesture is never totally excluded from the sound wave. If it were, the telephone would be a poorer instrument than it is.” (p. 205)

This functional gestural stance also underlies the function–form modelling in this paper. Prosody in speech interaction is regarded as the phonetic manifestation par excellence to jointly code information-related (propositional), listener-related (appellative), and speaker-related (expressive) meaning at any given moment, in different proportions, depending on the communicative function in the particular situational context [Kohler 2009]. A model of prosody that does not incorporate a communicative framework in its own right beside the formal level, to relate linguistic form and phonetic substance to communicative functions in all three semantic fields, is not capable of providing a comprehensive, descriptively and explanatorily adequate account of prosodic phenomena. AM Phonology does not fulfil this requirement. Furthermore, in view of multiparametric phonetic exponents in intricate function–form relations, discrete contrasts between prosodic units, with clearly marked boundaries between them, are the exception in speech communication, not the norm as postulated in AM phonology. This also casts
doubt on the hypothesis of strict segmental anchoring in the alignment of tonal targets with segmental boundaries, irrespective of rate or syllable structure [Ladd, Faulkner, R., Faulkner, H., Schepman 1999]. For instance, an $F_0$-peak or an $F_0$-valley contour-shift continuum across pre-accented, accented and post-accented syllables may, but need not, result in categorical perception in the Haskins sense of strong acuity across and weak acuity between category boundaries; yet the allocation to functional categories is always straightforward (see 1.2). There is category perception in all instances, but only some special cases are discrete categorical, i.e. have well-defined, narrow-band boundaries. Prosodic modelling needs to take these aspects into account, over and above the global communicative orientation. The Kiel Intonation Model (KIM) [Kohler 2006b] provides this type of function–form framework.

1.2 The Kiel Intonation Model (KIM)

KIM is based on the following postulates.

(a) $F_0$ patterns, like all other prosodic parameters, are related to categories of communicative function.

(b) Prosodic exponents identify communicative categories, without presupposing discrete boundaries for category differentiation.

(c) The short-time window of segment and the long-time window of prosody generation are synchronized in the coding of speech production and in the decoding of speech perception, with broad margins of segmental alignment around landmarks like accented vowel onsets.

(d) $F_0$-peak and valley patterns bear a holistic relationship to communicative categories and are not decomposed into local L and H tones of pitch accents, phrase accents and boundary tones at a formal prosodic phonology level.

(e) Although $F_0$ is certainly the most powerful prosodic marker of communicative categories, other parameters, e.g. energy and long-term phonation [Niebuhr 2010], timing of articulation and articulatory prosodies [Kohler 2011a,b; Niebuhr and Kohler 2011], also in conjunction with gestures, such as smiling [Kohler 2008], are further important contributors and need to be included in the analysis.

(f) The analysis of $F_0$-peak and valley synchronization has to include the shape of contours as a further factor in the identification of communicative functions [Dombrowski and Niebuhr 2005, Niebuhr 2007].
With regard to (b) and (c), it has been shown for German [Kohler 1987, 2005, Niebuhr and Kohler 2004] that in $F_0$-peak contour shifts the transition of the peak point into the accented vowel (from an *early* to a *medial* position) is perceived as a categorical change and allocated to a functional category change from **FINALITY** – ‘closing an argumentation’, to **OPENNESS** – ‘opening an argumentation’. The continuation of the shift towards the end of the accented vowel (to a *late* position) is, however, perceived as a gradual change to the functional category of **UNEXPECTEDNESS** – ‘contrast and expressive evaluation’. A comparable low-rising $F_0$-valley contour shift produces no categorical perception, but there is nevertheless a category change from **CASUALNESS** to **FRIENDLY CONCERN** when the low valley point moves from an *early* position before the accented vowel onset to a *late* position in the accented vowel. The same function–form relations apply to English (cf. Kleber [2006]), and no doubt to all West Germanic languages.

The synchronizations and shapes of peak and valley contours, according to (c) and (f), form holistic bundles in category coding, such that, for example in German, an earlier $F_0$-peak position combined with a slower fall from the peak maximum, after a faster rise to it, still signals a peak type for coding the **OPENNESS** category, and the opposite combination of synchronization and contour shape holds for a peak type coding **FINALITY**. This illustrates the importance of pitch relations in the $F_0$ movement relative to accented-vowel onset: in German (and other West Germanic languages) **FINALITY** is coded by *high-low*, **OPENNESS** by *low-high*. If $F_0$ stays high longer after a faster rise to a (higher) peak maximum, this strengthens high pitch in the first part of the accented vowel, even if the peak maximum is synchronized earlier, and this higher pitch is linked to the speech function **OPENNESS**. If $F_0$ falls faster after a slower rise to a (lower) peak maximum, low pitch is strengthened in the first part of the vowel even if the peak maximum is synchronized later, and the lower pitch is linked to the speech function **FINALITY**. This is another aspect of the **Frequency Code** [Ohala 1983].

Synchronization, contour shape, and maximum height of $F_0$ patterns are thus not independent factors in the production and perception of communicative functions; it is their convergence in generating different pitch dynamics around the accented vowel that codes the different functions. The communicative goal allows speakers a good deal of leeway in the timing of $F_0$ contours relative to articulatory landmarks, to adapt to the situational and contextual constraints in communication and to show individual preferences for synchronization or shape. Only the functional approach,
which enables the grouping of diverse measurement data to fixed categories in communicative behaviour that lie outside the analysed data, can give this insight.

In a preliminary experimental investigation, Niebuhr, D’Imperio, Gili Fivela and Cangemi [2011] compared synchronization and shape of $F_0$-peak contours in the Neapolitan Italian DECLARATIVE/INTERROGATIVE functions with the data found for FINALITY/OPENNESS in German. In the INTERROGATIVE function, the peak maximum is synchronized late in the accented syllable, whereas the DECLARATIVE function has early placement. These synchronization differences are accompanied by shape differences, but there was a subset of speakers who made very little use of synchronization and coded the opposition by greater divergence in shape. So, we find the same variability in the generation of a pitch-dynamics contrast among speakers of the two languages, but in Neapolitan Italian the *high-low vs. low-high* pitch pattern codes DECLARATIVE/INTERROGATIVE, for which the frequency code was originally developed.

1.3 Karl Bühler’s *Organon Model*

With its function–form relationship, KIM is part of a paradigm of speech communication which moves the expression of the speaker and the appeal to the listener beside propositional meaning into the centre of linguistic investigation. Its theoretical point of departure is Karl Bühler’s *Organon Model* [Bühler 1934]. This model links the linguistic sign to the Sender, the Receiver, and the Factual World of objects and factual relationships. The threefold association establishes the functions of EXPRESSION, APPEAL, and REPRESENTATION, by symptoms, signals, and symbols, respectively, as integral components in speech communication. They determine behavioural interaction in varying proportion and weight in any speech event at any given moment. Thus linguistics and paralinguistics are integrated into a comprehensive field of speech communication. The Speaker’s EXPRESSION comprises attitudes towards the Listener and the Factual World in communicative settings, emphatic evaluations, and emotions; APPEAL to the Listener is carried by commands, requests and questions; Factual World REPRESENTATION features in propositions and information structure.

1.4 Outline of the paper

The time is ripe for linguistics to start developing a comprehensive framework of communicative functions for prosody research. A fair amount of empirical work has already been done [Bolinger
1978, 1986, 1989; Ward and Hirschberg 1985; Kohler 2004, 2005, 2006a,b, 2009; Dombrowski and Niebuhr 2005; Niebuhr 2007, 2010; Niebuhr and Kohler 2004]. The goal of this paper is to develop four communicative functions on the basis of the Organon Model and of the differential weighting of its EXPRESSION symptoms, its APPEAL signals, and its REPRESENTATION symbols in any communicative speech event: Information Selection and Weighting (Section 2.1), Argumentation (Section 2.2), Declarative and Interrogative (Section 2.3), Intensification (Section 2.4).

Each functional field is introduced against a state-of-the art background, defined, together with its ramifications, and illustrated with reference to data from the two formally and structurally related languages English and German. The Interrogative Functions are focused on in fine detail, first in English (Section 2.3.1), then in German (Section 2.3.2). Subsequently, the developed framework of communicative functions is applied to data collection and analysis in the formally and structurally quite different tone language Mandarin Chinese (Section 3.1.1: Argumentation Functions, Section 3.1.2: Interrogative Functions, Section 3.1.4: Intensification Functions), and to the comparison of interrogative/declarative manifestations in Mandarin Chinese versus those in English and German (Section 3.1.3). Finally, the prosodic manifestations of interrogativity are related to the language-independent Frequency Code of homo loquens (Section 3.2).

Although Declarative and Interrogative are generally used with reference to syntactic structures, they refer here to an array of communicative functions, which are formally coded by syntactic structures and various prosodic devices. Traditionally, Germanic languages provided the word-order syntactic patterns for distinguishing between polarity questions and statements, and led to looking for differentiation by prosody if the syntactic pattern stayed declarative, first in the Germanic languages themselves, but then this formal perspective was extrapolated to languages like Mandarin Chinese as well, which do not have the different word-order structures. This prosody orientation on declarative syntactic structure paid little or no attention to different types of communicative functions within interrogativity. In this paper, the perspective is reversed, and fine functional differentiations are examined as to their formal manifestations.

The speech examples from the three languages are illustrated with graphic signal displays and audio files in wav format, indexed as Figure $n$ and Audio $n$ (click on Audio to listen).
2 Towards a Framework of Communicative Functions

2.1 Information Selection and Weighting

In verbal interaction communicators select information points, to which they give weight in relation to each other in the communicative situation. The languages of the world use a great variety of formal means to achieve this, e.g. structural devices, such as deviation from default declarative word order, cleft sentence constructions, lexical intensifiers, but also sentence accentuation by segmental lengthening (of vowel nuclei and/or initial consonants), and by pitch and energy patterns in specific syllables to make them salient for information highlighting of words and syntactic elements in declarative, interrogative and imperative functions. Traditionally this information selection and weighting has been studied in propositional meaning. It led to the categories of Focus and Presupposition or New and Given in a framework of information structure. Focus came to be studied in just one formal feature, namely $F_0$ patterns of pitch accents, which was initiated by the analysis of English [Liberman and Pierrehumbert 1984], but was also applied to German and Dutch.

In these West Germanic languages, focus can be freely marked by pitch. This does, however, not apply to a language like French.

In a contextualization of the famous English example from Liberman and Pierrehumbert [1984], we may get the following dialogue interchange:

Speaker A: Peter invited a few of his friends to a party in his flat. Mary came with Manny.

Speaker B: No, Anna came with Manny.

A has double focus and perceptually equal peak contour accents on Mary and Manny in typical sentence declination, (Fig. 1). B has single focus with a peak contour accent on Anna; Manny is either completely deaccented if the speaker only wants to set Anna in contrast to Mary (Fig. 2a), or it is partially deaccented with a low-rising valley contour if B picks up A’s background information as relevant for the statement (Fig. 2b).
Figure 1: Speech wave, spectrogram, $F_0$ (plain) and energy (dotted) of double focus peak patterns in *Mary came with Manny*. Male speaker, Southern British English (SBE). \[Audio1.\]

Figure 2: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) of single focus peak pattern on *Anna* in *No, Anna came with Manny*. \textbf{a} complete deaccentuation of *Manny*, \textbf{b} partial deaccentuation with late low-rising valley pattern on *Manny*; female speaker, SBE. \[Audio2.\]

Prosody-oriented studies of focus have been extended to a large number of languages, including Mandarin Chinese, in declarative and interrogative contexts [Xu 1999, Liu and Xu 2005, Liu 2009,
Jia and Li 2010, Jia, Li and Xiong 2010, Jia 2012]. They have been concerned with propositional meaning in information structure. Jia [2012], for example, collected data on numbers and places of focus in single sentences, elicited through *wh*-questions of the type “Who did what to whom before dawn?”, which were to be answered with sentences like Liu2 Min2 Ling2 Chen2 Ti2 Ba2 Mao2 Lan2 Le0. “Liumin promoted Maolan before dawn.” Both question and answer were presented in Chinese characters on a computer screen. This data acquisition procedure is not only far removed from communicative interaction between speakers in real situations, the above target sentence is also semantically odd. Wentao GU, Nanjing [personal communication] suggests that “even if the special thing happened before dawn, we would rather say Jin1 Tian1 Ling2 Chen2 Liu2 Min2 Ti2 Ba2 Mao2 Lan2 Le0 “Today, before dawn, Liumin promoted Maolan.” (which would, of course, destroy the formal uniformity of lexical tone throughout the sentence). The situational strangeness of meaning and the communicatively unusual multiple *wh* information questions are determined by a formal, rather than a functional, approach, and create a lab speech environment for prosodic data analysis whose relevance for information structuring in natural discourse remains unclear.

This paper moves in the opposite direction, from function to form, and differentiates between information and argumentation structure (see also the more natural data acquisition on focus in Gu et al [2006]). Argumentation structure is developed by communicators in ongoing discourse, parallel to, and different from, propositional information, introducing the speaker’s attitudes towards the listener and the factual world. In the examples of Fig. 2, for instance, a conveys contrastive focus in the sense that speaker B’s information contrasts with that of Speaker A in a factual way. But the comparison of a and b shows that in the latter, attitudinal argumentative contrast is added by later synchronization of a higher peak (see 2.2).

2.2 Argumentation

The speaker develops information points, selected for an utterance, into an argumentation structure with reference to four categories:

**FI** **FINALITY**

In a concluding argument, the speaker sums up his/her own apperception of a communicative outcome as being final and no longer debatable.

**OP** **OPENNESS**

In an opening argument, the speaker indicates that s/he has observed, and become aware of, something requiring further (inter)action.
UC UNEXPECTEDNESS – CONTRAST

In an opening argument, the speaker indicates with an overlay of contrast that observation contradicts expectation.

UE UNEXPECTEDNESS – EXPRESSIVE EVALUATION

In an opening argument with an overlay of contrast, the speaker adds expressive evaluation to an unexpected observation.

In English and other West Germanic languages, these argumentation categories are coded by differently synchronized pitch peak patterns (with possible ToBI notations): FI early (H+L*), OP medial (H*), UC late-medial (L+H*), UE late (L*+H), coupled with different maxima of $F_0$ and acoustic energy in comparison with UC. Fig. 3 illustrates OP, UC, UE in the English sentence He used to be slim., collected in the scenario of two people looking at old photos, one of them saying Here is an old photo of Ken., and the other commenting on Ken’s former body size. For further details, and illustrations of the functional categories in the corresponding German utterances Er war mal schlank. in the same situational contexts, see Kohler [2006b,2009].

Figure 3: Spectrograms, $F_0$ (plain) and energy (dotted) of a medial (top), b late-medial (centre) and c late (bottom) peak patterns in He used to be slim; a,c male speaker, b female speaker, SBE. [Audio3]
2.3 Declarative and Interrogative

The primary reference of the Declarative function is the factual world constructed through the speaker’s Argumentation. The main formal devices are syntactic structure and ellipsis, e.g. in response to questions, as well as pitch and other prosodic patterning. In non-tone languages, default utterance pitch tends to end low or falling in categorical statements. Concern for the listener intervenes as the speaker qualifies the validity of a statement by phrases of the type *I think*, *I suppose*, *may be*, question tags, or by final low rising pitch in English and German, or by voice quality (breathy voice).

On the other hand, the primary reference of the Interrogative function is the listener to whom the speaker makes an APPEAL to respond verbally, but there are several subcategorizations, again depending on varying APPEAL – EXPRESSION – REPRESENTATION constellations:

**IQ** INFORMATION QUESTIONS ask for specific information referring to subject, object, place, time, modality

- **IQ-1** with FACT ORIENTATION: matter-of-fact appeal for information
- **IQ-1.1** EXPRESSION OF CONTRAST
- **IQ-2** with LISTENER ORIENTATION: request appeal for information, ATTITUDE OF FRIENDLINESS
- **IQ-2.1** EXPRESSION OF CONTRAST

**PQ** POLARITY QUESTIONS ask for a decision as to the truth value of a proposition or argumentation along a positive-negative polarity scale

- **PQ-1** with LISTENER ORIENTATION: decision left open for the listener to make
- **PQ-1.1** EXPRESSION OF SURPRISE
- **PQ-2** introducing SPEAKER ORIENTATION: prejudging the listener’s decision towards one pole
- **PQ-2.1** EXPRESSION OF CONTRAST

**RQ** REPEAT QUESTIONS are an appeal to the dialogue partner for repeated verbal action: they pick up a dialogue partner’s preceding utterance and ask for an information point or the truth value to be confirmed

- **RQ-1** INFORMATION POINT ENQUIRY
- **RQ-1.1** EXPRESSION OF SURPRISE
RQ-2  TRUTH VALUE ENQUIRY

RQ-2a  REQUESTING CONFIRMATION of a dialogue partner’s assertion
RQ-2a.1  EXPRESSION OF SURPRISE

RQ-2b  EXPECTING CONFIRMATION of a dialogue partner’s previous assertion
RQ-2b.1  EXPRESSION OF SURPRISE

Whereas IQ is coded uniformly across the languages of the world by question words, e.g. English who, what, where, when, why, how, the coding of PQ relies on a great variety of formal means, including syntactic structure (initial verb position in the West Germanic languages), question particles (e.g. in Mandarin Chinese [Liu and Xu 2005, Liu 2009] and in literary Russian), or prosodic patterns on declarative syntax (e.g. in Neapolitan Italian [d’Imperio 2000]). But irrespective of the coding of the polarity function by syntactic or lexical means in a language, prosody codes subcategorizations. This prosody effect also applies to information questions. In RQ, prosodic patterns are of prime importance in combination with a variety of formal structures ranging from question-word to declarative syntax to elliptic phrases. The functional network of interrogativity will now be illustrated with reference to its formal coding in English.

2.3.1  Interrogative functions in English

In all three basic types of interrogative functions prosodic patterns combine with formal syntactic structures to code the various interrogative categories.

2.3.1.1  Category IQ: INFORMATION QUESTIONS

Question-word structure is combined with nuclear peak patterns for FACT ORIENTATION (IQ-1), but with nuclear valley patterns for LISTENER ORIENTATION (IQ-2). The peak contour has medial synchronization with the accented vowel (cf. 2.2; ToBI: H*L-L%). For the EXPRESSION OF CONTRAST (IQ-1.1), it has late-medial synchronization (cf. 2.2; ToBI: L+H*L-L%). The valley contour is low-rising (or falling-rising) and has late synchronization with the accented vowel (ToBI: (H+)L*H-H%). If the valley contour has early synchronization and is high-rising (ToBI: H*H-H%) the question-word structure no longer conveys an INFORMATION QUESTION but is a REPEAT QUESTION (RQ-1). The 4 different functional types of lexical questions are illustrated by Where? in the following dialogue and Fig. 4.
**Speaker A:** We’ll meet in Auchterarder tomorrow. (small town in Scotland, not widely known)

**Speaker B:** Where?

**IQ-1** With a *medial peak* B asks for more information about the location of the venue in the town. This introduces the functional meaning of OPENING ARGUMENT (cf. 2.2) into the question context. (Fig. 4a)

**IQ-1.1** With a *late-medial peak* B stresses the need for more information about the venue against the insufficiency of the information so far given by A. This superimposes the functional meaning of CONTRAST on the OPENING ARGUMENT (cf. 2.2), and introduces it into the question context. The utterance has a tone of irritation and impatience: “But where? Your information is rather imprecise.” (Fig. 4b)

**IQ-2** With a *late low (falling-)rising valley*, where the rise starts in the accented vowel, B still asks for more information about the venue, but makes a request appeal to the listener. The fall adds contrast (**IQ-2.1**), but the utterance sounds less categorical and more friendly than with a peak pattern. (Fig. 4c)

With an *early high-rising valley*, where the rise starts before the accented vowel, Speaker B appeals to Speaker A to repeat the name of the place because s/he has not heard properly or finds it strange (interrogative function **RQ-1**, Fig. 4d, cf. 2.3.1.3 REPEAT QUESTIONS).

---

**Figure 4:** Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) of 4 interrogative functions in the question-word structure *Where?* a *medial peak* in **IQ-1**, b *late-medial peak* in **IQ-1.1**, c *late valley* (low falling-rising) **IQ-2**, d *early high-rising valley* in **RQ-1**; female speaker, SBE. [Audio 4]
2.3.1.2 Category **PQ**: POLARITY QUESTIONS

Interrogative word-order structure is combined with nuclear valley patterns in LISTENER ORIENTATION (PQ-1), but with nuclear peak patterns in SPEAKER ORIENTATION (PQ-2). The valley contour is high-rising and has early synchronization with the accented vowel (ToBI: H*H-H%). For the EXPRESSION OF SURPRISE (PQ-1.1), the high rise has late synchronization (ToBI: L*+HH-H%). The peak contour PQ-2 has medial synchronization with the accented vowel (cf. 2.2; ToBI: H*L-L%). For the EXPRESSION OF CONTRAST (PQ-2.1), it has late-medial synchronization (cf. 2.2; ToBI: L+H*L-L%). The 4 different question types are illustrated by *Is he in Rome?* in the following dialogue and Fig. 5.

**Speaker A1**: Where is he?

**Speaker B**: He has gone to Italy.

**Speaker A2**: Is he in Rome?

**PQ-1** With an *early high-rising valley* where the rise starts before the accented vowel, A does not prejudge the answer but appeals to the listener for a polarity decision. (Fig. 5a)

**PQ-1.1** With a *late high-rising valley*, where the rise starts in the accented vowel, A still appeals to the listener for a polarity decision, but with an expression of surprise at the person perhaps being in Rome. (Fig. 5b)

**PQ-2** With a *medial peak*, A wants more information about the person’s whereabouts and suggests a place, expecting the answer to be *yes*. This introduces the function of OPENING ARGUMENT of a medial peak contour into the question context. (Fig. 5c)

**PQ-2.1** With a *late-medial peak*, A wants more information, as in PQ-2, but contrasts his/her suggestion with his/her expectation. (Fig. 5d) This introduces the superimposed functions of CONTRAST and OPENING ARGUMENT of a late-medial peak into the question context.
Figure 5: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) of 4 interrogative functions in the polarity-question structure *Is he in Rome?* a *early valley* in PQ-1, b *late-valley* in PQ-1.1, c *medial peak* in PQ-2, d *late-medial peak* in PQ-2.1; male speaker; SBE.

[Audio5]

The functional explanation of the use of rising or falling pitch patterns in PQ-1 and PQ-2 polarity questions is buttressed with data presented by Fries [1964]. In his extensive American English corpus (39 television-radio programmes in which a panel of four persons, using, in turn, only yes-no questions, attempted to discover the precise vocation, occupation, or special activity of each of several contestants), he found, over all speakers, 61.7% examples with falling and 38.3% with
rising intonation. This result was contrary to the textbook statement given for English. He comments on it as follows: “The circumstances in which the programmes were carried on made the speech forms used by these panellists the actual live conversation of language actively fulfilling its communicative function. The speed and spontaneity of the language activity of these panellists reduced to practical zero the chance that the intonation forms of that language activity could have been premeditated or deliberately chosen.” (p. 247)

In the structuralist tradition, Fries only provides the empirical data without attempting to explain them. He concludes: “The facts seem to support the conclusion that in English (at least in American English) there is no question intonation pattern as such. ... when one compares the intonation patterns of all yes-no questions with the intonation patterns of all other types of questions, he will find that, even with the ratio 3 to 2 in favour of falling intonation patterns for yes-no questions, which the evidence here supports, there will be a higher proportion of rising intonation patterns on yes-no questions than on other questions. But there seem to be no intonation sequences on questions as a whole that are not also found on other types of utterances, and no intonation sequences on other types of utterances that are not found on questions.”(pp. 250f)

Fries leaves unexplained why his data show the opposite trend to what the textbooks say. The key to an insightful explanation is in the communicative situation: panellists ask their questions so as to ascertain the highest possible number of yes responses in order to win the game, and thus prejudge positive answers by proposing their own constructions of facts, i.e. the addressee’s free decision between the answers yes and no is not a prime concern. This is the typical case of speaker and fact orientation, and therefore associated with falling pitch. It does, of course, not mean that listener orientation does not occur; it is common when questions are asked again because they were not heard clearly or not understood by the contestants. In this case, it was the same question, asked a second time by the same panellist and directed to the same person, almost immediately after the first one. In this repetition, listener orientation comes in after the factual question has been asked, i.e. rising pitch is very likely. There are also cases where the first utterance of such a question pair has rising intonation, signalling listener orientation, and the second has falling pitch turning factual. This sequence is also found in cases such as

Speaker A1: Are you coming with us? (as an APPEAL for a decision, with rising intonation)
Speaker B: What? (as an APPEAL to repeat, with early high-rising intonation, cf. 2.3.1.3)
**Speaker A2:** Are you coming with us? (no longer as an APPEAL to the dialogue partner but as a factual REPRESENTATION of what has already been asked, with falling intonation).

Although Fries is right in rejecting a “question intonation” as such, determined solely by the syntactic form, the probability of rising pitch in word-order questions is higher than in question-word questions and vice versa for falling pitch. The two semantic-pragmatic question types IQ and PQ establish different relations between the speaker and the factual world, on the one hand, and the listener, on the other. In an IQ, the speaker prototypically asks about the factual world, in a PQ, the speaker prototypically appeals to the listener. In English, these prototypically different question functions are coded in prototypically different forms of question-word questions with falling pitch and word-order questions with rising pitch. In specific situations, speakers restructure this link by reducing its prototypical orientation and strengthening its complement, according to the communicative demands: an information enquiry becomes friendly, personalized with rising pitch, a polarity decision gets prejudged with falling pitch. The actual pitch manifestation of questions in dialogues therefore depends on the interaction of general pragmatic function and attitudinal colouring in specific communicative situations. The use of pitch in questions is thus directly related to semantic-pragmatic function and needs to be explained on the basis of some general behavioural principle that triggers it. The orientation towards the listener by the wide-spread use of high or rising $F_0$ in polarity questions in the languages of the world will be discussed in 3.2 with reference to such an external principle, the Frequency Code.

### 2.3.1.3 Category RQ: REPEAT QUESTIONS

In INFORMATION POINT ENQUIRY, REPEAT QUESTIONS **RQ-1** occur with question-word structure and a high-rising valley pattern. It starts on the question word and has early synchronization, see Fig. 6a (cf. also Fig. 4d). Late synchronization adds the EXPRESSION OF SURPRISE (**RQ-1.1**). The latter may be further heightened by breathy phonation, increased duration, increased $F_0$ rise and increased energy for NEGATIVE INTENSIFICATION, see Fig. 6b, cf. 2.4. The dialogue in 2.3.1.1

**Speaker A:** We’ll meet in Auchterarder tomorrow.

**Speaker B:** Where?

provides examples.
Figure 6: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) of 2 repeat questions in the question-word structure *Where?* a *early high-rising valley* in RQ-1, b *late high-rising valley* in RQ-1.1 – expression of surprise with negative intensification; female speaker, SBE. [Audio6]

In **Truth Value Enquiry Requesting Confirmation** of a dialogue partner’s assertion (RQ-2a), a high-rising valley pattern is used in declarative or elliptic syntax, e.g. in the dialogue

**Speaker A**: He has gone to Rome.

**Speaker B**: He is in Rome?! To Rome?

It has early synchronization, Fig. 7a, or late synchronization for the additional **Expression of Surprise** (RQ-2a.1), Fig. 7b. The latter may again be heightened for **Negative Intensification** (RQ-2a.1-NI), Fig. 7c, cf. 2.4.
The request for confirmation may be strengthened by the addition of a constant-polarity question tag in the rising pitch movement of the repeated assertion: “He is in Rome, is he?” The tag “reinforces verbally the questioning intonation” [O’Connor, 1955, p. 102]. Neither a peak pattern on the tag nor a reversed polarity of the tag would be possible in this REPEAT QUESTION context. “The [constant-polarity] tag question refers back to a fact already and recently established by the listener, whereas it is precisely the listener’s view the reversed tag question seeks to elicit .... This difference of background explains why it is impossible with the [constant-polarity] tag question to have the pitch pattern of fall plus fall except with a violent disjunction.” [O’Connor 1955, p. 102] “...a falling tone would demand agreement from the listener, but demanding agreement when the listener has already himself presented the information is pragmatically inappropriate.” [Cruttenden 1997, p. 98]

The reversed-polarity tag does not refer back to the dialogue partner’s preceding statement, but is a POLARITY QUESTION, in which the speaker makes an assertion of his/her own in a new statement, and then asks for confirmation in a question tag. This question either prejudges agreement, with falls on both the declarative structure and the tag [O’Connor 1955, p. 98], or expresses increasing
doubt as to the truth value of the declarative, with a low or high rise on the tag [O’Connor 1955, p. 99], thus leaving the answer open in “a norm al conducive question” [Bolinger 1986, p. 389]. The functions of POLARITY QUESTIONS with declarative structure + question tag or with interrogative structure differ, requesting either confirmation or a polarity decision on the speaker’s assertion. In both cases, the speaker prejudges the answer or leaves it open by using falling or rising intonation.

If a dialogue partner’s previous assertion is picked up with the contextual meaning of EXPECTING CONFIRMATION, declarative or elliptic syntax is used with a peak pattern in a high register. Thus in *He is in Rome?/To Rome?*, the high register signals the question, the declarative syntax the repeat appeal, and the peak pattern points to the expected confirmation. These formal devices signal that confirmation of the truth value of what has been said previously is taken for granted, and that an answer is optional: **RQ-2b**. This is opposed to *He is in Rome?/To Rome?* with an early high-rising valley pattern, which requests confirmation of a queried truth value and solicits an answer: **RQ-2a**.

The use of a high-register peak pattern with a declarative structure for EXPECTING CONFIRMATION in a REPEAT QUESTION is not well documented in the literature on English intonation. In the practical orientation of the London School descriptions of intonation, the communicative function of this REPEAT QUESTION form was of minor importance for teaching and learning English as a foreign language. On the other hand, experimental data acquisition methods need to use very sophisticated dialogue contextualizations to elicit these function–form patterns, and AM laboratory phonology has not developed such a methodology. But there are a couple of references in Bolinger [1986].

“She /said she would /never be /happy with /him.

Once again, the speaker might utter all this in a monotone – as a sort of musing question, at a fairly high pitch level but with little or no variation in pitch.” (p. 17)

“For example, an utterance like

\[\text{N} \hat{\text{e}} \text{body be}
\]

\[\text{l} \text{eved it, then?}\]

in which the pitch goes down to signify a conclusion of sorts, but (and that is the gestural counterpart of the question mark – … the eyebrows go up – the speaker manages to assert and ask at the same time, which is taken to mean ‘assertion presented for confirmation’.” (p. 204)
The pitch marking indicates a wide range with high-level strengthening.
Cruttenden [1997, p. 84] refers to Repeat Questions as follows.
“Echoes are most commonly questions which query the whole or some part of the previous utterance of another speaker, often with a note of incredulity, e.g.
(I didn’t go after all.) You didn’t GO?
(Take two hours this morning to get that overseas order ready.) Two HOURS?
(What about going to Ascot tomorrow?) Go to ASCOT?
Echoes can be exclamatory rather than questioning, e.g.
(He’s got a distinction.) A DISTINCTION!
(Get that bit of wood for me.) Get that bit of WOOD! Just who do you think you’re talking to?
A difference is often made between an echo question and an echo exclamation by the choice of tone: echo questions commonly take high-rise, echo exclamations commonly take rise-fall.”
The first three examples could also have peak patterns on a high register and still be Repeat Questions with Confirmation Expected (RQ-2b), where an answer is optional. This gets Cruttenden’s echo exclamations close to echo questions.

2.3.1.4 Sequential restrictions on questions in dialogue
Since RQ-2 Repeat Questions take up a dialogue partner’s pronouncement as a whole, and enquire its truth value, they cannot ask for a polarity decision on new, additional aspects connected to it. So, “He is in Rome?” would not follow “He has gone to Italy.”, unless the speaker takes it for granted that the person would be in Rome when in Italy. The Polarity Question “Is he in Rome?”, types PQ-1, PQ-1.1, is the usual enquiry in this situational context.

Contrariwise, since a Listener-Oriented Polarity Question (PQ-1) asks for a yes-no decision it cannot be used when the answer is already in a dialogue partner’s previous assertion, and the speaker wants to have its truth value confirmed. So, Is he in Rome? with an early high-rising valley will not follow He has gone to Rome. But with a peak pattern it can, as this no longer solicits an open decision from the dialogue partner, and refers to the already given statement as the answer, with a note of unexpectedness. The surprise may either concern the locality in contrast to other places, in which case Rome will have a late peak, and the prehead will be low. Or it may be concerned with being in Rome, in the sense Is that really so?, in which case the whole utterance is in a high register, indicating that an answer, if at all solicited, will be
expected to confirm what has already been stated: the interrogative structure is now a REPEAT QUESTION. *Is he in Rome?* with a late, instead of an early, valley for the EXPRESSION OF SURPRISE (PQ-1.1) is quite possible after the statement *He has gone to Rome*, also with NEGATIVE INTENSIFICATION (cf. 2.4). In the same context, the declarative structure *He is in Rome?* is combined with an early high-rising valley pattern to *ask* for confirmation, or with a peak pattern in a high register to *expect* confirmation (cf. 2.3.1.3).

But if *He is in Rome.* has a peak pattern in a non-high register it can no longer be a REPEAT QUESTION and therefore cannot follow a dialogue partner’s statement *He has gone to Rome.* because restating what has already been stated would need some introduction, such as *aha, I see*, or a following constant-polarity question tag with a low rise after a falling pattern on the statement *He is in Rome, is he?* (cf. 2.3.1.3). Similarly, *He is in Rome.* with an early or late low-rising valley pattern is a statement with listener orientation, and does not become a REPEAT QUESTION, so is excluded from the statement context *He has gone to Rome.*, unless a constant-polarity tag question follows (*He is in Rome, is he?*), continuing the rise of the statement (cf. 2.3.1.3). Both types of question tags add an APPEAL to confirm the statement the speaker picked up from the dialogue partner. So, the strengthening of high pitch by raised register in peak patterns or by high $F_0$ end points in valley patterns is crucial for the signalling of REPEAT QUESTIONS.

### 2.3.2 Interrogative functions in German

German uses the same bundling of the same types of formal means as English to signal the same sub-categorizations of the interrogative function:

- four syntactic structures – lexical, interrogative, declarative, elliptic
- two phrase intonations – rising and falling
- a raised pitch level in declarative or elliptic structures with falling intonation, to distinguish question from assertion.

According to Kohler [1977, p. 199, 1995, p. 197], the four utterances

- *Ist er gekommen?* “Has he arrived?” – rising intonation
- *Ist er gekommen?* “Has he arrived?” – falling intonation
- *Er ist gekommen?* “He’s arrived?” – rising intonation
• *Er ist also gekommen?* “He’s arrived, has he?” – falling intonation are all semantic questions, but with very different functions, which are realised with different formal devices, showing that questions are by no means always associated with rising “question intonation”, as text books generally maintain (Niebuhr et al. [2010]). In the fourth example, falling intonation is even combined with declarative structure for expected confirmation. This REPEAT QUESTION function is strengthened by the modal particle *also*, which corresponds to the use of a constant-polarity low-rising question tag after a fall on the statement: *He has come, has he?* (cf.Kohler [1978]). The 2x2 syntax–prosody forms are referable to the functions of INFORMATION, POLARITY and REPEAT QUESTIONS and their sub-categorizations, within the framework developed in 2.3. The functional differentiations of the two pitch patterns in question-word questions and in word-order questions have been analysed on the basis of German corpus data in Kohler [2004, 2009]. The following sections present the data for the interrogative functions in parallel with the English data.

2.3.2.1 Category **IQ**: INFORMATION QUESTIONS

**IQ**

*Speaker B*: *Wo?* “Where?” in the dialogue context of

*Speaker A*: *Das Treffen findet in Hademarschen statt.*

“The meeting will take place in Hademarschen.” (a small town in the north, not widely known in Germany).

**IQ-1**

With a *medial peak*, B asks for more information about the location of the venue in the town. (Fig. 8a)

**IQ-1.1**

With a *late-medial peak*, B stresses the need for more information about the venue against the insufficiency of the information so far given by A. The utterance has a tone of irritation and impatience. (Fig. 8b)

**IQ-2**

With a *late high-rising valley*, where the rise starts in the accented vowel, B still asks for more information about the venue, but makes a request appeal to the listener. The utterance sounds less categorical and more friendly than with a peak pattern. (Fig. 8c)

With an *early high-rising valley*, where the rise starts before the accented vowel, Speaker B appeals to the listener to repeat the name of the place because s/he has not heard it properly or finds it strange (interrogative function **RQ-1**, Fig. 8d, cf. 2.3.2.3 REPEAT QUESTIONS).
2.3.2.2 Category PQ: POLARITY QUESTIONS

PQ  Speaker A2:  *Ist er in Rom?* “Is he in Rome?” in the dialogue context of
Speaker A1:  *Wo ist er denn eigentlich?* “I wonder where he is.”
Speaker B:  *Er ist nach Italien gefahren.* “He has gone to Italy.”

PQ-1  With an *early high-rising valley*, where the rise starts before the accented vowel, A
does not prejudge the answer but appeals to the listener for a polarity decision. (Fig. 9a)

PQ-1.1  With a *late high-rising valley*, where the rise starts in the accented vowel, A still
appeals to the listener for a polarity decision, but this time with an expression of
surprise at the person perhaps being in Rome. (Fig. 9b)

PQ-2  With a *medial peak*, A wants more information about the person’s whereabouts and
suggests a place, expecting the answer to be *yes*. (Fig. 9c)

PQ-2.1  With a *late-medial peak*, A wants more information, as in PQ-2, but contrasts his/her
suggestion with his/her expectation. (Fig. 9d)
Figure 9: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) of 4 interrogative functions in the polarity-question structure *Ist er in Rom?:* a early valley in PQ-1, b late-valley in PQ-1.1, c medial peak in PQ-2, d late-medial peak in PQ-2.1, male speaker; Standard German [Audio9].

2.3.2.3 Category RQ: REPEAT QUESTIONS
The category of REPEAT QUESTIONS with elliptic syntax is illustrated by the following conversation I overheard in a café between a young woman and a young man. The man said that he had moved and was now able to get to work more quickly.
*Woman:* Mit deinem Auto, oder wie? “In your car, or how?”

First part falling on raised level, tag rising, requesting confirmation


*Rising,* querying previous statement and requesting confirmation

*Woman:* Ehrlich? “Really?”

Man: Ja, bei jedem Wetter. “Yes, whatever the weather.”

*Woman:* Ehrlich? “Really?”

Falling on raised pitch level, complimentary question, confirmation expected.

*Man:* No further comment on this subject

The English Repeat Questions in 2.3.1.3 have the following German equivalents:

**RQ-1** INFORMATION POINT ENQUIRY

*Speaker B:* Wo? in the dialogue context of

*Speaker A:* Das Treffen findet in Hademarschen statt.

With question-word structure and high-rising valley pattern, starting on the question word, in early or late synchronization, depending on the absence or addition of an

**EXPRESSION OF SURPRISE** RQ-1.1. (Fig. 10a,b)

![Figure 10: Speech waves, spectrograms, F0 (plain) and energy (dotted) of 2 repeat questions in the question-word structure Wo? a early high-rising valley in RQ-1, b late high-rising valley in RQ-1.1 – expression of surprise with negative intensification; male speaker, Standard German. Audio10.](image)

28
RQ-2  **TRUTH VALUE ENQUIRY**

*Speaker B:* Er ist in Rom?/Nach Rom? “He’s in Rome?”/“To Rome?” in the context of

*Speaker A:* Er ist nach Rom gefahren. “He’s gone to Rome.”

**RQ-2a**  **REQUESTING CONFIRMATION** of a dialogue partner’s assertion,
with declarative or elliptic syntax and a high-rising valley pattern

- in early synchronization for a matter-of-fact **REPEAT QUESTION** (Fig. 11a)
- in late synchronization for additional **EXPRESSION OF SURPRISE RQ-2a.1** (Fig. 11c)
- potentially heightened for **NEGATIVE INTENSIFICATION RQ-2a.1-NI**, cf. 2.4 (Fig.11d)

**RQ-2b**  **EXPECTING CONFIRMATION** of a dialogue partner’s previous assertion

with declarative or elliptic syntax and a peak pattern in high register (Fig. 11b).

**REPEAT QUESTION RQ-2a** requests confirmation of a queried truth value and solicits an answer,
**REPEAT QUESTION RQ-2b** takes confirmation of the truth value of what has been said previously for granted, and considers an answer optional.
Figure 11: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) of 4 repeat questions in the declarative structure Er ist in Rom? a early high-rising valley in RQ-2a, b raised-register peak in RQ-2b, c late high-rising valley in RQ-2a.1 – expression of surprise, d late high-rising valley in RQ-2a.1-NI – expression of surprise with negative intensification; male speaker, Standard German. [Audio11]

A typical matter-of-fact statement, on the other hand, e.g. in the context

**Speaker A:** Wo ist er denn eigentlich? “I wonder where he is.”

**Speaker B:** Er ist in Rom./In Rom. “He is in Rome.”/“In Rome.”

has a peak pattern on a non-raised pitch level with speaker or fact orientation, or a low-rising early or late valley pattern with listener orientation. The early or late valley expresses absence or presence of interest to give an answer: CASUALNESS vs. FRIENDLY CONCERN (Fig. 12a-c, [Audio12]). They are not possible as declarative-structure REPEAT QUESTIONS in the context Er ist nach Rom gefahren., cf. [Audio12_1]. If the statement conveys the expression of agitation or even irritation (“Everybody knows that and so should you!”) the peak pattern is raised. It can be identical with the one in a repeat question RQ-2b, where confirmation is expected, although the stated fact came as a surprise, cf. [Audio12_2]. In both cases, the higher pitch level is triggered by increased activation of speech production to stimulate a listener to give an answer, or to express tense attitudes. If the statement is made with the expression of surprise at the other person asking (“Have you forgotten that he told us some time ago?”) the same late high-rise pattern can be used as in the surprise repeat question of RQ-2a.1, cf. [Audio12_3].
Figure 12: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) of 3 statements in the
Er ist in Rom. a non-raised medial peak, b early low-rising valley, c late low-rising valley,
Standard German. [Audio12]

These uses of the same prosodic forms for different functions in different verbal and situational
contexts show that questions and statements cannot be defined by formal features but need to be
referred to functional-semantic categories, which together with the communicative setting of
EXPRESSION, APPEAL, and REPRESENTATION determine their phonetic (prosodic and segmental) as
well as gestural exponency. Bolinger [1978, pp. 503f] highlights this function–form relationship
and situation dependency:
“...It is not by any means certain that “question” is a grammatical category at all to the extent that it is
marked only by intonation. And if it is not, then we need to ask what the intonations mean,
independently of any grammatical type. An utterance such as

And he believes that.

...can be a question, a statement, or an exclamation, depending on context and gesture. But the
intonation is just as conclusive one way or another. As a question the sentence is inquisitive, it
probably calls for confirmation of what is already assumed. With fuller descriptions, we may
find the same variety prevailing everywhere.” (cf. also Bolinger [1989], chapters 5. Questions,
6. Nonquestions)
2.4 Intensification

Niebuhr [2010] distinguishes 3 types of INTENSIFICATION (cf. also Kohler [2006a]),

NI  NEGATIVE INTENSIFICATION: disapproving expressive evaluation of a statement
PI  POSITIVE INTENSIFICATION: praising expressive evaluation of a statement
RI  REINFORCEMENT: emphatic underscoring of the validity of a statement

with the following prosodic exponents in German declaratives. In NI, initial consonants of accented syllables are considerably lengthened, phonation is tense breathy, $F_0$ forms a pointed peak quite early in the vowel, with a fast falling slope. In PI, the accented vowels are considerably lengthened, phonation is breathy-voice, $F_0$ rises into a plateau followed by a shallow fall. In RI, consonant lengthening and $F_0$ peak shape are as for NI, but synchronized later; phonation is not tense breathy. NI and PI also occur in the interrogative function [Kohler 2011a], in both cases with late valley synchronization, but tense breathy in NI, breathy-voice in PI.

The same function–form relations apply to English. Thus, “Where?” as an INFORMATION QUESTION with the EXPRESSION OF CONTRAST (IQ-1.1) may be given NEGATIVE INTENSIFICATION with falling pitch. It expresses exasperation over not getting precise information. Is he in Rome? as a POLARITY QUESTION with an EXPRESSION OF SURPRISE (PQ-1.1) may also be given NEGATIVE INTENSIFICATION, this time with a late valley contour. The same applies to REPEAT QUESTIONS He is in Rome? (RQ-2a .1) and Where? (RQ-1.1), all expressing incredulity.

Fig. 13 compares In a handbag. as a simple Repeat Question RQ, a Repeat Question with the Expression of Surprise and Negative Intensification NI, “Lady Bracknell style”, and a Repeat Question with the Expression of Surprise and Positive Intensification PI (cf. Kohler [2011b]).
Figure 13: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) traces of *In a handbag*. a RQ b NI c PI; female speaker, SBE.

3 Applying the Framework in Comparative Prosodic Research

The framework of communicative functions outlined in Section 2 is a theoretical construct of human interaction, based on observation in European languages. By extrapolation it may be taken as a basic postulate for speech communication in any language. Linguistic analysis then needs to find out how the functions are formally manifested in the languages of the world, thus to advance comparative prosodic research by giving it new direction. In a first step, some preliminary Mandarin Chinese data have been collected and analysed within this communicative framework, and are set against the German and English manifestations.

3.1 Application to Mandarin Chinese

Since tonal features are tied up in the lexical tones of a tone language, it is a prime question in comparative prosodic research how speakers implement the categories of the functional framework as overlays of the lexical distinctions. The following discussion of a small set of data from Mandarin Chinese attempts preliminary answers.

3.1.1 Argumentation Functions

As regards the Argumentation functions, the extensive data analysis of focus in declarative and interrogative structures in isolated sentences of Mandarin Chinese [Xu 1999, Liu and Xu 2005, Liu
is not sufficient to gain insight into communicative, as against linguistic, functions. It has only dealt with information structure in propositional meaning and cannot give answers about argumentation structure. Some data have been presented within the communicative perspective for hao-3, xing-2 “OK” by Kohler [2009] (recorded by Yi XU, UC London - Fig. 14, and Aoju CHEN MPI Nijmegen - Fig. 15). As suggested by Aoju CHEN, the assertions may be contextualized as follows:

**Finality:** the speaker closes the argument and may express reluctance and even resignation

Your boss asks you to hand in a project proposal soon. You explain that this is not feasible for various reasons. But your boss insists that your company needs the proposal quickly.

*Your boss:* “Will you then hand in the proposal in two weeks?”

*You:* hao-3, xing-2 (Figs. 14a,c, 15a,c)

**Openness:** the speaker is open to the argument and may express willingness to comply

Your boss asks you to hand in a project proposal soon. Because it is a very short proposal and you have a clear idea of what it should be like, you think it can be done quite easily.

*Your boss:* “Will you hand in the proposal in two weeks?”

*You:* hao-3, xing-2 (Figs. 14b,d, 15b,d).

Figure 14: Speech waves, spectrograms, $F_0$ (plain) and energy (dotted) traces of Mandarin Chinese hao-3 a FI b OP and xing-2 c FI d OP; male speaker, [Audio14](#).
For **Finality (FI)** the low tone ends low and concomitantly acoustic energy is low or decreases, whereas for **Openness (OP)**, pitch rises and energy is high or increases. Similarly, the high tone ends lower or higher and concomitantly the time course of acoustic energy is on a lower or higher level. In the **FI** context, “OK” in English is realized as an *early peak*: $F_0$ falls rapidly in the accented second vowel to a low level from high in the preceding unstressed vowel. In the **OP** context, it has a *medial peak*: $F_0$ rises into the accented vowel from a medial level of the unstressed vowel, and then falls. **FI** strengthens low pitch, **OP** high pitch in the accented syllable (see 2.2). So, in Mandarin Chinese pitch lowering or raising is at work to code the argumentation functions of **Finality** and **Openness** as it is in English and German, simply adapted to the conditions set by the tone language.

### 3.1.2 Interrogative Functions

Questions have been studied extensively, but always from the angle of linguistic form. Even the detailed analysis of parallel encoding of interrogative meaning, focus and lexical tone in Mandarin Chinese by Liu and Xu [2005] and Liu [2009] takes formal question types as the point of departure. Subjects read isolated sentences, graphically marked, among others, as statement by a period, *ma* particle question or so-called yes-no question, i.e. in declarative syntax, with a question mark. This leaves out the essential situational embedding of different types of interrogative function in speech communication. Moreover, asking each speaker to read 380 sentences out of context cannot
guarantee a reliable separation of the functional question types, although regularities of the influence on lexical tone patterns have been clearly demonstrated.

Fig. 16 provides instances of statements, *ma* and declarative-structure questions, which I collected from a female native speaker (Xiaojun ZHAO, 38yrs), reading the 4 tonal frames in Liu [2009] from Chinese script, marked by final period or question mark. In each tonal sequence, the register is raised for both question types across the whole sentence vis-à-vis the statement. In the rising sequence, the last rising tone additionally goes up very high, but this high rise is curtailed by the final neutral-tone *ma*. In the low-tone sequence, the last low tone falls and ends in creak in the statement, but rises high in the questions, also before the neutral-tone *ma*. In the sequence of falling tones, the last one ends at a higher level in the questions as well, which is particularly striking when compared with the tone at the internal phrase boundary before a pause.
Zhang Wei dànxīn XiǎoYīng kāichē fāyūn
ZhangWei worry XiaoYing driving dizzy
“Zhang Wei worries that XiaoYing will get dizzy while driving”
b  (Tone 2 Rising) [Audio16 2]

Wáng Méi huái yí LiúNíng huáchuán zhāomi
WangMei suspect LiuNing canoeing obsessed

“WangMei suspects that LiuNing will get obsessed with canoeing”
“LiMin dislikes LiuYu to light a fire to keep warm” (Liu’s citation tone transcription is kept, although sandhi transforms tone 3 in the first syllable of each word to tone 2.)
Figure 16: Spectrograms and $F_0$ traces, in each of the 4 frames of a high, b rising, c low and d falling tone sequences: in each panel, statement (top), declarative-structure question (centre), ma particle question (bottom); a transcription window is linked to the latter; female speaker.

In all cases, the questions have a higher articulation rate than the statements. After an internal phrase break, there is tonal reset, at the statement or question register level, which is more marked in the questions. So, it is not just the declarative-structure question that has higher pitch marking, spread across the whole utterance, to express interrogativity, but this applies to the ma particle question as well. However, these data do not give any information about the functional difference
between the two question forms, nor about the interpretation the speakers gave to the orthographic form when asked to read the communicatively strange sentences out of plausible context.

In the next step therefore, questions in Mandarin Chinese need to be investigated in a communicative framework to find out what the formal realizations of the different Interrogative functions, superimposed on lexical tone, are and how they compare with the function–form relationship found in the West Germanic languages. To this end, I collected some preliminary data from the same speaker by devising short English dialogues between a husband and wife who are wondering how to occupy their boy. The contextualization gave a high probability of eliciting POLARITY and REPEAT questions, with and without SURPRISE, as well as statements including CONTRAST. The script was translated into Mandarin Chinese by the speaker and her French husband Alexis Michaud, CNRS. Some of the dialogues are presented below; the target utterances are underlined.

**Husband:**

wǒmen jīntiān dài èrzi qù nàr wán ne? nǐ zěnme xiǎng?
“What shall we do with the boy today? What do you think?”

A tā xīhuàn qù dōngwùyuán/hàitān ma?
A “Does he want to go to the zoo/beach?”

**Wife:**

wǒ xiǎng tā bù yuànyì qù. B Tā xīhuàn qù hàitān/dōngwùyuán.
“I don’t think so.” B “He wants to go to the beach/zoo.”

**Husband:**

C repeat questions requesting confirmation

C1 tā xīhuàn qù hàitān/dōngwùyuán?
C1 “He wants to go to the beach/zoo?”

C2 tā bùshì gèng xīhuàn qù dōngwùyuán/hàitān ma?
C2 “He does not prefer the zoo/beach?”

**D** repeat question with disbelieving surprise

D tā xīhuàn qù hàitān/dōngwùyuán? zhēn nán yīxiāngxin!
D “He wants to go to the beach/zoo? I can hardly believe that.”
The dialogues were read by the speaker in both roles of wife and husband. Future systematic data collection will have to record proper dialogues with pairs of a male and a female speaker and also swap the texts in the two roles so that complete data sets will be obtained from all female and male speakers. In the dialogue embedding, $\text{A}$ is a polarity question of the open, listener-oriented type (cf. PQ-1 in 2.3), according to Wentao GU, on the strength of Audio 17. $\text{B}$ is a statement, which may also be contrastive ($\text{B-c}$) (cf. OP and UC in 2.2); $\text{C1}$ is a repeat question requesting confirmation (cf. RQ-2 in 2.3); $\text{C2}$ reinforces the confirmation request of the preceding repeat question (cf. RI in 2.4); $\text{D}$ is a repeat question with disbelieving surprise (cf. RQ-2a.1 in 2.3).

Figs. 17 and 18 compare non-emphatic and emphatic functions, respectively: $\text{ma}$ questions $\text{A}$ (polarity) and $\text{C2}$ (reinforced repeat), statements without ($\text{B}$) and with contrast ($\text{B-c}$), prosody questions $\text{C1}$ (repeat) and $\text{D}$ (surprise repeat) in $\text{tā}$ (bùshì gèng) xīhuan qù hāitān (ma). Figs. 19 and 20 do the same for $\text{tā}$ (bùshì gèng) xīhuan qù dòngwùyuán (ma).
Figure 17: Spectrograms, $F_0$ (plain) and energy (dotted) + transcription for Tā xīhuàn qù hǎitān (ma): ma polarity question A (top), statement B (centre), prosody repeat question C1 (bottom); female speaker. [Audio17]
Figure 18: Spectrograms, $F_0$ (plain) and energy (dotted) + transcription for Tā (bùshì gèng) xīhuan qù hǎitān (ma): reinforced ma repeat question C2 (top), contrastive statement B-c (centre), prosody surprise repeat question D (bottom); female speaker, Audio18.
Figure 19: Spectrograms, $F_0$ (plain) and energy (dotted) + transcription for tā xīhuan qù dòngwùyuán (ma): ma polarity question A (top), statement B (centre), prosody repeat question C1 (bottom): female speaker, Audio19.
Figure 20: Spectrograms, $F_0$ (plain) and energy (dotted) + transcription for tā (búshì gèng) xǐhuan qù dòngwùyuán (ma): reinforced ma repeat question C2 (top), contrastive statement B-c (centre), prosody surprise repeat question D (bottom); female speaker, [Audio20.]

The ma POLARITY QUESTION A differs from the statement B by both syllables of hǎitān being raised, and $F_0$ then descending slightly for the neutral-tone ma; in dòngwùyuán the low point of the fall and the high point of the rise are raised, $F_0$ levelling out in ma. These patterns differ from the ma questions in the corresponding frames of Fig. 16. Although the syllable tones are shifted upwards both in Fig. 16 and in A of Figs. 17/19, compared with the statements, the difference is concentrated on the utterance-final word, hǎitān or dòngwùyuán, in the latter, but distributed more
generally over the whole sentence in the former. Wentao GU has suggested the following explanation for this difference. “The statement B in the dialogue is immediately preceded by wǒ xiǎng tā bú yuàn yì qù ("I don’t think so.")”, which makes the statement B different from a context-free isolated read sentence, introducing a certain degree of contrastiveness on the sentence-final word hǎitān or dōngwùyuán (especially on the latter, where it is clearly perceptible in Audio 19) – though the degree of contrastiveness is smaller than in the statement with an explicit CONTRAST (B-c). This contrastiveness lowers the $F_0$ minimum of dōngwùyuán, and hence a larger upward-shift is observed for the ma polarity question.”

**Prosody** **REPEAT QUESTIONS C1** differ from statements B by having the entire tone sequences shifted upwards in both sentences. Moreover, final high-tone tān rises instead of being level, and final rising-tone yuán rises continuously very high instead of evening out at a lower level. These data tie in with the statements and the prosody questions of the 4 tonal frames in Fig. 16. In the **prosody SURPRISE REPEAT QUESTION D**, the falling-rising $F_0$ contour of low+high-tone hǎitān is raised, and the $F_0$ contour of falling+rising-tone dōngwùyuán is expanded in its low turning and its high end point, compared with the repeat question C1. In addition, the final syllable in both cases of D ends in breathiness. In the **ma REPEAT QUESTION C2**, the speaker reinforces his request for confirmation C1 by negating the object of his polarity question A. Compared with the polarity question, hǎitān ma in the repeat question is shifted up under focus, and the slight dome shape contour is greatly expanded; dōngwùyuán has its fall and its rise expanded.

When the statement is made by contrasting the communicatively focused word of B with the one in question A, the final syllable of hǎitān or dōngwùyuán is given greater prominence than in the non-contrastive statement, by a 20-30% increase of its duration, by increased acoustic energy, and by tense phonation. Contrastive hǎitān also has a full final nasal consonant with descending $F_0$ resulting in a dome-shaped contour for the high syllable tone, which is level in the non-contrastive case.

The examples indicate that a raised register feature, superimposed on the lexical tone sequence of a declarative phrase structure, signals a repeat question. Expanding the $F_0$ contour of the focused word and overlaying it with breathiness adds the meaning component of disbelieving surprise. Increasing the prominence of the focused word adds contrastiveness to the statement. Finally, all the
emphatic functions in Figs. 18 and 20 have a higher level of acoustic energy compared with the non-emphatic ones in Figs. 17 and 19.

3.1.3 Comparing Interrogative vs. Declarative Functions in Chinese with English and German

In Mandarin Chinese as well as in English and German, raised pitch plays an important role in the signalling of Repeat Questions. In Mandarin Chinese, declarative syntax combines with high pitch register across the entire question, compared with the statement. This also occurs in English and German for Truth Value Enquiry Expecting Confirmation, but for Requesting Confirmation the main pitch effect resides in the final high rise starting on the last accent. When surprise is added to the request for confirmation, English and German use a late, instead of an early, high-rising valley pattern, Mandarin Chinese expands the F0 contour of the tones in the focused word. In both Mandarin and the West Germanic languages, the surprise component may be further intensified, among others, by breathiness (cf. 3.1.4).

Polarity questions in English and German are word-order questions with either a nuclear fall or rise, depending on whether the speaker prejudges an affirmative answer or leaves the decision entirely to the listener. In Mandarin Chinese, polarity questions are ma particle questions. The collected data point to two different sentence intonation patterns: either the pitch register of the entire sentence is raised, compared with the corresponding declarative-syntax statement, or raising is concentrated on the pitch contours of the tones in the focused word. Whether this pitch extension difference reflects the same functional difference that is coded by pitch direction in English is a question to be solved by further data acquisition and analysis.

With C2, the speaker renews his request for confirmation immediately after his first repeat question C1, by negating the boy’s preference for the other locality. He may do this either by insisting on querying the truth value, or by requesting a response as to whether the negation is true. In the former case, the prosody question form would simply be repeated on the negated statement structure. In the latter case, which was realized in the present data collection, the Mandarin speaker uses a ma particle question with reinforcement of the focused word. In English or German, the same function may be coded by a high register shift of an entire nuclear-fall pitch pattern (see 2.3.1.3,
2.3.2.3), with the possible addition, in English, of the tag question doesn’t he? as a post-nuclear low rise.

The contrastive statement B-c is realized by similar means in both Mandarin Chinese and the West Germanic languages to increase the prominence of the focused word. But apart from accented-syllable lengthening, the $F_0$ peak contour is also expanded in English and German.

### 3.1.4 Intensification Functions

The prosody surprise repeat questions D and the reinforced ma repeat questions C2 exemplify negative intensification and reinforcement in questions, as illustrated for English and German in 2.4. In spite of the large-scale study of focus in Mandarin Chinese, e.g. Liu and Xu [2005], Liu [2009], the systematic analysis of the Intensification functions under focus has been largely excluded. Linguists across the world’s languages have shied away from investigating this function because belonging primarily to attitudinal and expressive meaning it has not been deemed to be a proper topic in formal linguistics and has therefore been relegated to paralinguistics as part of the study of emotion. However, this is the field where communicative functions are very likely coded, to a large extent, by the same prosodic means in the languages of the world. The use of breathiness in the Mandarin Chinese examples D of Figs. 18, 20 and in the intensified English and German repeat question of Figs. 6b, 7c, 13b and 10b, 11d bear witness to this.

### 3.2 The Frequency Code

Since higher pitch, either higher register or rising, serves to differentiate questions from statements in both Mandarin Chinese and English, adapted to the presence or absence of lexical tone constraints, it is legitimate to ask whether we are dealing here with a universal feature in all languages. Ohala [1983] answered in the affirmative by proposing the Frequency Code, which means that in asking a question a speaker subordinates to a listener, and subordination is coded by a universal mechanism of raising pitch in speech. This version of the frequency code can certainly be regarded as an ingredient in coding and decoding questions. But when the subcategorization of questions is taken into account it becomes doubtful that Ohala’s Frequency Code can provide a general explanation for all types of questioning.
The function of REPEAT QUESTIONS, without or with surprise, is to attract attention and stimulate a dialogue partner into action, rather than a passive subordination to the partner’s response. And POLARITY QUESTIONS in which the speaker prejudgethe answer tend to have falling pitch in West Germanic languages. Furthermore, closer inspection of the interrogative function in speech communication has demonstrated the need for a subtle differentiation of many situationally determined types of questions which it is impossible to subsume under the same production principle.

It was, for instance, shown for German in Kohler [2011b] that the precursor to an accentual high rise may also be raised to a higher level. In a perception experiment using the Semantic Differential technique, listeners judged repeat questions with high precursors as expressing a more agreeable attitude towards the addressee (on a contrary – agreeable scale) than repeat questions with low precursors. The same may be expected for English. As the high rise in the repeat question has an activation function, this stimulation is intensified if the difference in pitch between the utterance beginning and the end of the rise is increased, i.e. by a low precursor, whereas the high precursor softens the repeat activation, making it more accommodating. In German and English, repeat questions may also have falling pitch patterns in declarative syntax transposed to a higher register to elicit confirmation that what is enquired IS true (see 2.3.1.3, 2.3.2.3 and 3.1.3). This differs from the high-rising repeat question, which enquires the truth value one way or the other.

These subtle differentiations of communicative question functions need a more sophisticated Interrogative Code, which goes beyond the Frequency Code in trying to explain more than the linguistically triggered broad dichotomy of question and statement. The Interrogative Code encapsulates the multifarious relations in communicative interaction between speaker and listener where the listener-oriented APPEAL function is central, but gets adjusted in various ways by the speaker’s attitudes and expressiveness. There seems to be overwhelming evidence that the APPEAL function is coupled with high pitch [Hermann 1942, Bolinger 1978], but the coexistence of both high rising and high register in the same language to code different question types makes it mandatory to develop a universally valid Interrogative Code cautiously within a framework of communicative functions.

Moreover, Rialland [2007] has provided data from African register tone languages of the Sudanic region that do not show the high-pitch link with questions. Various combinations of an open vowel
question marker, sonorant segment lengthening, delayed falling intonation and breathy termination occur to signal a question on the same syntactic structure as a statement. She refers to these property bundles as a ‘lax prosody’. However, it is not clear what position these questions have in the framework of interrogative functions that has been presented in this paper. Most examples are translated as statements vs. repeat questions, e.g. beans. beans?, a slave. a slave? It is to be expected that these languages also have the functional categories of polarity and repeat questions, and use different prosodic means to signal them.

4 Conclusion and Outlook

On the basis of function-oriented data acquisition and prosodic analysis of argumentation structures, questions, and emphasis in English and German it has been possible to propose a communicative framework of Information Selection and Weighting, Argumentation, Declarative and Interrogative, and Intensification functions, which control the formal exponents of speech interaction in these languages. The functions have been regarded as basic in human communication with reference to Bühler’s Organon Model, allowing their extrapolation to other languages as a general communicative frame into which each language fits its own formal means for coding the inter-language functions. This “function first, form second” approach, in turn, makes it possible to use the language-independent framework of communicative functions as a powerful tool in comparative prosodic research. However, its application necessitates a fundamental change in data acquisition methodology, because reading off isolated sentences will no longer do; tightly contextualized utterances in plausible texts are needed that can capture the functions to be investigated.

In a first step, the functional framework has been applied to the analysis of some formal exponents of Argumentation, Declarative and Interrogative, and Intensification in Mandarin Chinese, focusing in greater detail on Declarative and Interrogative. Data collection was based on situationally contextualized dialogues and contrasted with data from the reading of isolated sentences of systematically selected syntactic structures and lexical tone sequences. But the acquisition procedure still lacked the proper simulation of interaction between dialogue partners. Subsequent investigation will have to refine the methodology of devising dialogue scripts and of implementing them in recording sessions with speaker pairs. Moreover, the three functional domains have not received equal coverage. The Argumentation categories Finality and Openness, with overlays of
CONTRAST and EXPRESSIVE EVALUATION, and the three types of Intensification still require a great deal of attention. Even the Interrogative functions have not been examined sufficiently, although they have been dealt with more thoroughly than the other two domains.

Furthermore, a lot more data need to be collected from many more speakers for systematic instrumental analysis and statistical evaluation of the prosodic parameters fundamental frequency, acoustic energy, articulatory timing and phonation in the implementation of these communicative functions by speakers, and in their relevance for listeners. Future analysis also needs to extend the proposed function–form investigation to a great variety of languages for insightful comparative prosodic research in speech communication. To provide an appropriate communicative basis for comparing varying formal devices, this cross-linguistic function–form analysis should, first of all, include those languages, e.g., Danish, French, Italian, Spanish, Swedish, where Focus and Sentence Mode have been studied quite extensively, mainly with isolated read sentence data (but see the functional approach to Swedish intonation in Ambrazaitis [2009]). The comparative study needs to include register tone languages of Africa as well, for which diverging question coding has been reported by Rialland [2007]. It will have to be ascertained what type of interrogative functions show this divergence and how the different types of interrogativity are manifested.

Such a cross-linguistic function–form approach will lead to a new, communicatively insightful prosodic typology of the world’s languages, and to a reassessment of traditionally postulated universal codes, like the Frequency Code, as generally valid aspects of prosody in the coding of specific speech functions. The theoretical framework of communicative functions itself needs to be further developed and refined. This paper is a programmatic proposal to stimulate such function-oriented prosodic research, to reinstate a goal Hermann [1942, p. 390] formulated 70 years ago for future integral linguistic research of the world’s languages (“ganzheitliche Sprachforschung”): capturing “the distribution of different rhythmic-melodic properties over the different speech functions and their summation in one and the same function.” (“die Verteilung der verschiedenen rhythmisch-melodischen Mittel auf die verschiedenen Funktionen und ihre Summierung bei ein- und denselben Funktionen”).

Acknowledgements
This paper is a revised and enlarged version of the plenary lecture “Developing a Framework of Communicative Functions for the Study of Speech Prosody”, which I was invited to give at the Third International Symposium on Tonal Aspects of Languages (TAL2012) in Nanjing. My special
thanks are due to Wentao GU for inviting me to a stimulating and very enjoyable conference, for
giving me much appreciated comments, particularly on the Chinese data, and for having the paper
reviewed by two anonymous reviewers, whose suggestions for revision were extremely helpful. I
am also very grateful to Aoju CHEN and Yi XU for making recordings for me, and to Fang LIU for
stimulating exchanges on the analysis of questions in Mandarin Chinese and in English, following
my reading of her PhD thesis. I am deeply indebted to Alexis Michaud and Xiaojun ZHAO for
insightful discussion of my English dialogue scripts, for translating them into Mandarin Chinese
and for recording them so competently. Any inadequacy in the interpretation of the data is my own.

References
de Lund 49 (2009).
1931).
Bolinger, D.: *Intonation across languages,* in: Greenberg (ed.), *Universals of Human Language,*
1989).
Ohio State University (2000).
Dombrowski, E.; Niebuhr, O.: *Acoustic patterns and communicative functions of phrase-final F0
Fries, C.C.: *On the intonation of ‘Yes-No’ questions in English.* In D. Abercrombie, D.B. Fry, P.A.D.
Longmans 1964).
Gu,W., Hirose, K., and Fujisaki, H.: *Modeling the effects of emphasis and question on fundamental
frequency contours of Cantonese utterances.* IEEE Trans. Audio, Speech and Language


