Intonational Universals and Intonational Typology

D. R. Ladd
Edinburgh University
December 1999

This is the final draft of chapter 98 ("Intonation"), prepared for *Language Typology and Language Universals: An International Handbook*, edited by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible. The chapter appeared, with minor corrections, in the published *Handbook* (Mouton de Gruyter, 2001), pp. 1380-1390. The present version is for informal use only; any citation should refer to the published version.

Contents

1. Introduction

2. Intonational Universals
    2.1. Substantive Universals
    2.2. Universalist Interpretations
    2.3. A Phonological Interpretation

3. Intonational Typology
    3.1. Melodic Typology
        3.1.1. Lexical vs. Intonational
        3.1.2. Intonational Melodic Features
    3.2. Accentual Typology
        3.2.1. Lexical vs. Intonational
        3.2.2. Cross-linguistic Differences of Sentence-stress
        3.2.3. Phonological or syntactic typology of sentence-stress?
    3.3. Prosodic Typology
        3.3.1. Syntax-Prosody Mapping
        3.3.2. Sentence-stress and Phrasing

4. Intonational Phonetics
    4.1. Declination
    4.2. Overall Pitch Range
    4.3.Dynamic vs. Melodic Accent

5. References

1. Introduction

Intonation occupies a somewhat unusual place in any discussion of language universals and language typology. In dealing with most aspects of language one takes for granted that languages differ: what typology is about - and in a sense what linguistics is about - is identifying the patterns of similarities amid the differences. But this basic assumption does not necessarily hold for intonation. A great deal of influential research - for example, the work of Bolinger (e.g. 1978, 1989), Lieberman (1967), and Fujisaki (e.g. 1983) - is based on what will be referred to here as the *universalist* approach to intonation. In this view, intonation is essentially peripheral to language. It reflects a pre-linguistic communicative code, and/or it depends on physical properties of the speech production mechanism. The important features of intonation are actually the same in all languages; any observable differences are merely conventionalized variations on the universal themes, and therefore any typological generalizations we may find will be different in kind from those that can be drawn about phonology or syntax. In recent years the universalist view has lost some ground to what we might call the *phonological* approach, based on the work of Bruce (1977), Pierrehumbert (1980), Beckman (1986) and others. As we shall see in §2.3, this approach treats intonation as part of phonology, and as such is compatible with the idea that normal typological generalizations about intonation may be possible. Nevertheless, the universalist view still thrives, and one important task in writing about intonational universals and intonational typology is to do justice to both points of view.

Section 2 of this article presents the undisputed cross-language similarities of intonation first (§2.1), together with a universalist interpretation (§2.2) and a critical evaluation of that interpretation (§2.3). Section 3 then summarizes three aspects of intonation – melody (§3.1), sentence-stress (§3.2), and phrasing (§3.3) – that differ across languages in a way that reflects possible typological dimensions. Finally, §4 briefly discusses several dimensions of what we might call the phonetic typology of intonation.

2. Intonational Universals

2.1. Substantive Universals

The most basic use of intonation is to divide the stream of speech into chunks, which we will refer to here as intonational phrases or IPs. (The term IP is used in a pretheoretical way, refer to intonationally defined subunits of utterances, and should not necessarily be identified with e.g. the intonation phrase of Nespor & Vogel 1986 or Pierrehumbert & Beckman 1988; see §3.2.3 for further discussion.) The existence of IPs - and hence IP boundaries - provides the basis for many of the cross-linguistic generalizations emphasized by universalists. These generalizations focus o overall pitch trends across IPs and with pitch movements at the ends of IPs. The most important observations are the following: (1) pitch tends to decline from the beginning of an IP to the end, a tendency known as declination (Pike 1945, Cohen & 't Hart 1967, Cooper & Sorensen 1981, Ladd 1984); (2) the beginning of an IP may be marked by a local sharp rise in pitch or "reset" (Brown, Currie & Kenworthy 1980, 't Hart, Collier & Cohen 1990); (3) in IPs that are utterance-final and/or in statements, there may be a local drop in pitch at the end of the IP in addition to any overall declination spanning the IP as a whole (Lieberman 1967, Liberman & Pierrehumbert 1984); (4) in IPs that are in questions and/or are not utterance-final, declination may be moderated, suspended or even reversed, i.e. the overall trend may be less steeply declining, level, or even slightly rising (Thorsen [Grønnum] 1980, Lindau 1986); (5) in addition to exhibiting reduced declination, non-

final and interrogative IPs may also have a local rise in pitch at the end, or at least have no local drop (Lieberman 1967). The validity of these observations, as general tendencies, is not in doubt.

The other basic use of intonation is to signal local prominence or emphasis. In a great many languages, particular syllables or words can be picked out from the stream of speech by some combination of increased loudness and/or duration, greater force and/or precision of articulation, and pitch movement and/or widened pitch range. In many languages these features are put to use in the unquestionably linguistic phenomenon of "stress" or "accent", but even in languages which do not seem to have stress or accent, and in languages where pitch is lexically specified ("tone languages" and "pitch accent languages") it is still possible to identify intonational cues to local prominence or emphasis. For example, in both Chinese (Gårding 1987) and Japanese (Pierrehumbert & Beckman 1988), pitch is lexically specified, but local emphasis can be signalled by a local widening of the overall pitch range.

2.2. Universalist interpretations

Two types of explanation have been offered for the substantive universals just sketched. The more common of the two, which we might call the biomechanical view, emphasizes the universal physical characteristics of the speech production mechanism. The mechanisms of speech obviously require speakers to stop talking every so often to breathe, which results naturally in the division of speech into IPs (indeed, Lieberman (1967) and others refer to IPs as "breath groups"). At the same time, the cognitive mechanisms of speech seem to operate in terms of meaningful constituents of some sort, and in all languages, so far as is known, there is a tendency for IP boundaries to correspond with the boundaries between such constituents. Universalists emphasize the naturalness of the correspondence between breath pauses and constituent boundaries (but see further §3.3). The physical mechanisms of speech also require the expenditure of breath, which may lower subglottal pressure, which in turn may manifest itself as lower fundamental frequency and hence declination across the IP. The release of subglottal pressure at the end of phonation naturally creates a sharp drop in pitch at the end of the IP. The combination of natural tendencies to declination and final pitch drop leads to a universal association between low pitch and finality or completeness. Given that association, an IP that ends without a pitch drop will sound incomplete; this incompleteness may then signal either that the speaker has not finished talking (high pitch for continuation), or that the speaker's utterance needs to be completed by an utterance of the addressee's (high pitch for interrogation). As for local prominence, local rises in pitch can be explained in terms of momentary increases in subglottal pressure, force of articulation, etc., which will perturb the overall trends across the IP but will not otherwise affect them. Explanations along these general lines have a long history (e.g. Lieberman 1967, Vaissière 1983).

Besides the biomechanical view, intonational universals have also been explained with reference to universal sound symbolism. This view is strongly identified with Bolinger's work (e.g. 1978, 1989) and has been developed by e.g. Cruttenden (1981). Rather than emphasizing the mechanisms of speech production, this view posits the existence of a universal, prelinguistic "meaning" of high and low pitch, part of a universal code for signalling states of the organism. High pitch signals interest, tension, incompleteness, etc., while low pitch signals resolution, completeness, and rest. This view is not necessarily incompatible with the biomechanical view, but at the very least it emphasizes different aspects of the problem. For example, in the sound symbolic view, the high pitch that begins an IP bears some semiotic relation to the high pitch that signals local emphasis, even though they may be produced by different mechanisms. Mention should also be made of the difficulty of describing sound symbolic meanings, which makes it difficult to evaluate competing hypotheses. For example, Ohala (1982) has also proposed that there is a universal pitch code - universal to all mammals, in fact - based on the metaphor of high pitch for smallness and submissiveness and low pitch for bigness and dominance. It is not clear whether Ohala's proposal is compatible with Bolinger's.

The universalist view, and in particular the biomechanical version of it, often goes hand in hand with what Ladd (1983, 1996) has called the *overlay* model of intonation. In the overlay view, contours result from the interaction of two components, which (following Fujisaki) we will refer to as the *phrase component* and the *accent component*. The details differ from one description to another, but in general the idea is that the phrase component gives an overall characterization of the trend of pitch across an IP - gradually declining, level, etc. - while the accent component (which can be extended to model lexical tone in languages like Chinese) generates local pitch movements that are superimposed on the phrase-level trend. The two broad types of intonational universal (IP-related and prominence-related) thus correspond to two distinct aspects of speech production. However, the overlay model is logically distinct from the universalist view, and some proponents of overlay models emphasize their language-specific application (e.g. Grønnum 1995, Möbius, Pätzold & Hess, 1993).

2.3. A phonological interpretation

To some extent, the question of whether intonation is basically universal (with minor variations) or basically language-specific (with many broad cross-language similarities) is not an empirical question but a matter of emphasis. However, one important argument against the universalist approach is that it is effectively unfalsifiable. A concrete example will make this clear. Under many circumstances, yes/no questions in many languages of Eastern Europe (e.g. Greek, Hungarian) end with a sharply rising-falling pitch movement spanning the last two or three syllables. This contour sounds to speakers of many Western European languages (e.g. English, Italian) like an emphatic *declarative* contour. That is, there is no simple cross-linguistic interpretability of pitch contours, which is what the strongest version of the universalist theory would lead us to expect. The universalist generalizations are broad enough to accommodate this kind of cross-linguistic fact: the Eastern European question contour can be treated as a combination of rising or high pitch for questions and a sentence-final fall that signals the end of the utterance, while the Western European emphatic declarative contour can be treated as an example of high pitch for emphasis and sentence-final fall for finality. But the very ability to accommodate this kind of cross-linguistic ambiguity severely limits the predictive power of the theory.

As noted in the introduction (§1), a good deal of recent work on intonation has replaced the universalist view - and the overlay model - with an explicitly phonological approach that Ladd (1996) refers to as the "autosegmental/metrical" theory. In this theory, a pitch contour in any language is treated as a sequence of tones, some of which are related to local prominences (or in the case of "tone languages", to individual lexical items), and others of which are related to IP boundaries. The inventory of possible tones, and the question of whether tones are lexically or intonationally specified, is part of the phonology of a language; the details of the association between tones and segmentals are governed by prosodic structures that are in part language-specific.

A central feature of the autosegmental/metrical approach is that the phonological elements of which pitch contours consist are *local* "events", not global shapes or slopes. Overall trends across IPs (e.g. gradual lowering of pitch) are mostly modelled as the result of repeated local effects (e.g. repeated downstepping of accents in succession). There is some instrumental evidence for this sequential view of overall trends (e.g. Liberman & Pierrehumbert 1984), and there is more general evidence for the view that IP edges may be locally marked by tones. Even in traditional works that assume the overlay view, it has been noted that in some tone languages (such as Chinese [Chang 1958] or Thai [Abramson 1962]) sentence-level intonational meanings are conveyed by modifications of the *last* lexical tone in a phrase or utterance. This is easily modelled as the local effect of an additional "boundary tone" associated with the final syllable.

3. Intonational typology

We now turn to the aspects of cross-linguistic intonational variation that appear to be of typological interest. There are three such aspects, which we refer to here as *melodic, accentual,* and *prosodic.*

3.1. Melodic typology

3.1.1. Lexical vs. intonational

As is well known, pitch features in "tone languages" are specified lexically. Such lexically assigned pitch features are not a part of intonation, and are not treated here (see article 123). However, the interaction of lexical and intonational features appears to be of typological interest. As we saw in §2.3., Chinese and Thai appear to have lexical tones on most syllables and, in addition, intonational boundary tones at the ends of IPs. However, as noted by Ladd (1996:149) there are some tone languages (e.g. Yoruba) that seem to have only lexical tones, and no boundary tones at all. Further combinations of lexical and intonational tones may be found in so-called "pitch accent" languages: in Tokyo Japanese, the accent tones are lexical but the boundary tones intonational (Pierrehumbert & Beckman 1988); in several European pitch accent languages it appears that some accent tones are lexical while other accent tones, like boundary tones, are intonational (Gussenhoven & Bruce 1999). The foregoing list is merely suggestive of the typological possibilities, which so far remain almost entirely unexplored.

3.1.2. Intonational melodic features

Concentrating only on languages without lexically specified pitch ("intonation languages"), we may observe melodic differences of several types. First, languages may use "the same" tune in different ways. For example, the English rising nuclear tune is widely used in both questions and statements in North America, Australia, and New Zealand, but is largely restricted to echo or confirmation questions in most British varieties (Cruttenden 1994, Ladd 1996). Second, languages may use different tunes for the same basic function (e.g. the differences in basic yes/no question intonation discussed in §2.3). Third, some languages (e.g. many European languages) seem to have a greater variety of possible tunes than others (e.g. Indonesian, Odé and van Heuven 1994). None of these differences lead to clear typological generalizations in our present limited state of knowledge. However, the phonological approach acknowledges the existence of the differences, and does not attempt to reduce them all to variations on the same universal pattern. Research is now being devoted to issues that had previously escaped notice (such as the precise alignment of pitch targets relative to the segmental string, e.g. Prieto, van Santen & Hirschberg 1995; Arvaniti, Ladd & Mennen 1998; Ladd, Faulkner, Faulkner & Schepman 1999). There is growing evidence of systematic cross-linguistic differences of intonation.

One such difference involves the treatment of *tonal crowding*. This arises particularly in the case of accented phrase-final monosyllables; in e.g. *Did she buy a car?,* the pitch on *car* may be required to rise and then fall and then rise again. In some languages (e.g. English) all the specified pitch movements in such cases are normally realized fully. In others there are various ways of dealing with tonal crowding that avoid multiple pitch movements on a single syllable. For example, in Greek, pitch movements may begin earlier and the pitch range may be compressed (Arvaniti, Ladd, and Mennen, 1998); in Hungarian, the pitch movement may simply be truncated (Ladd 1996). The typological distinction between "truncation" and "compression" is explored by Grønnum (1991).

3.2. Accentual typology

3.2.1. Lexical vs. intonational

As with melodic features, it is necessary to distinguish between word-level (lexical) accentual properties and sentence-level (intonational) ones. Word-level accent, or word-stress, has always been a topic of importance within generative phonology and is the subject of a large literature. Some of this literature is explicitly typological, e.g. Halle & Vergnaud's attempt (1987) at a principles-and-parameters description of possible word-stress systems (see also Hayes 1995, and for a rather different approach Guierre 1979). Word-stress is treated separately in this handbook (article 122) and is not discussed further here. The rest of this section deals with sentence-level accent, or sentence-stress.

3.2.2. Cross-linguistic differences of sentence-stress

We begin with the universalist position that sentence-stress, as part of intonation, is essentially the same in all languages. In this view, individual words are rendered prominent – made to stand out phonetically in some way from their context – in order to "highlight" them, to signal their importance or salience in the discourse. This view is espoused by Bolinger (e.g. 1978, 1986) and by quite a number of other broadly "functionalist" researchers. If we accept this view, then there can be no meaningful typology of sentence-stress.

However, a number of recent works have identified ways in which sentence-stress appears to differ cross-linguistically. A case in point involves *deaccenting:* Ladd (1990, 1996), Vallduví (1992), Cruttenden (1993), and Zubizarreta (1998) have all presented evidence that languages systematically differ in the extent to which sentence-stress occurs on repeated or contextually given words (i.e. the extent to which repeated or contextually given words are "deaccented"). For example, in English the repeated *have* would be deaccented as in (1a), and the form with both occurrences of *have* accented (1b) is very odd. Exactly the reverse is true in the structurally parallel Italian sentences (2a) and (2b).

(1) a. Better to HAVE one than NOT to have one.
b. * Better to HAVE one than not to HAVE one.

(2) a.* Meglio AVERne che NON averne.
b Meglio AVERne che non AVERne.

On the basis of such differences, Vallduví (1992) proposes the existence of a typological parameter [plastic]: in [+plastic] languages like English, the accentual pattern can be molded to express information structure, whereas in [-plastic] languages like Catalan or Italian, the accentual pattern is fixed and the job of expressing information structure falls to syntax. Related ideas are explored by Zubizarreta (1998), although her formal analysis is rather different.

Another apparent difference between languages involves sentence-stress in questions. Ladd (1990, 1996) has drawn attention to the fact that in some languages, including English, French, and German, sentence-stress in questions does not differ from that in statements: to the extent that the most prominent word is the rightmost content word in statements, the same is true in questions, both yes/no questions and question-word (WH) questions. In other languages, including Greek, Hungarian, and Russian, questions and statements differ: yes/no questions have sentence stress on the finite verb, while WH-questions have sentence stress on the WH-word. This can be seen in the contrast between the English examples in (3) and the closely parallel Greek equivalents in (4).

(3) *Statement*: He bought CHOCOLATES.
Yes/No question: Did he buy CHOCOLATES?
WH-question: Where are the CHOCOLATES?

(4) *Statement:* Agorase SOKOLATES.
Yes/no question: AGORASE sokolates?
WH-question: PU ine i sokolates?

Ladd suggests that, cross-linguistically, questions may either exhibit the sentence stress pattern of statements (as in English), or else consistently diverge from the statement pattern (as in Greek), but it may also be that sentence-stress patterns in yes/no questions and WH-questions vary independently. In any case this is a rich potential field of typological research. . The reader is referred to Ladd (1996) for more discussion of various aspects of sentence-stress typology.

3.2.3. Phonological or syntactic typology of sentence-stress?

The cross-linguistic differences just discussed are very much *sui generis:* there is no obvious way in which a tendency to deaccenting can be related to other aspects of a language's phonology, nor any obvious link between sentence-stress patterns and the syntax of questions. In the long run this is surely unsatisfactory. If there are typological patterns to sentence-stress, it seems likely that they will be part of a larger typological picture. However, attempts to relate sentence-stress to other aspects of phonology or syntax are still very preliminary.

One such attempt is the work of Halle & Vergnaud (1987), who explicitly treat the location of sentence-stress (prominence within the IP) as a phonological matter, like the location of word-stress (prominence within the foot or phonological word). In Halle & Vergnaud's terms, IPs, like feet or phonological words, are either "right-headed" or "left-headed". (The most prominent element within a prosodic domain is its phonological "head".) The parameters of variation posited to account for word-stress (left-headed vs. right-headed, bounded feet vs. unbounded feet, etc.) apply at the level of sentence-stress as well. Among other things, this view predicts that some languages typically have prosodic prominence on the last (rightmost) word of an IP while others have it on the first (leftmost).

On the face of it, Halle & Vergnaud's claim seems implausible, because, as Cinque (1993) points out, it predicts typological variety in sentence-stress systems of a sort that does not appear to be attested. Concretely, in many languages the greatest prominence seems to occur at the right end of the IP. This was the essence of Chomsky & Halle's Nuclear Stress Rule for English (1968), and Cinque himself espouses a modified version of this claim as valid for all languages. (Specifically, Cinque says that sentence-stress goes on the most deeply-embedded terminal node of the syntactic tree; superficial differences between languages reflect syntactic differences, not sentence-stress differences as such.). Yet a prediction very much like Halle & Vergnaud's is made, on rather different grounds, by Nespor and Vogel (1986), who claim that relative prominence within the IP (their "phonological phrase") is a function of syntactic headedness. Specifically, they state that syntactically head-final languages are phonologically left-headed (i.e. have greatest prominence at the beginning), while syntactically head-initial languages are phonologically right-headed (i.e. have greatest prominence at the end). They provide a variety of examples that seem to support their claim.

It is beyond the scope of this article – and beyond the scope of our present knowledge – to evaluate these competing views. However, it is worth noting that they may be partially reconciled by assuming that there are actually two different levels of IP. Many proposals have been made over the years distinguishing what we might call "major IPs" from "minor IPs", with a major IP consisting of one or more minor IPs. The currently most influential version of this idea is Beckman

& Pierrehumbert's (1986) distinction between *intonation phrases* (major) and *intermediate phrases* (minor). It seems accurate to state that Nespor & Vogel, and perhaps also Halle & Vergnaud, are primarily concerned with differences at the level of the minor IP, which correspond roughly to syntactic "maximal projections" (noun phrases, prepositional phrases, etc.). Cinque's claims, on the other hand, apply mainly to major IPs, which tend to correspond to whole sentences and parenthetical phrases, and which probably do tend to have their greatest prominence on the right.

Examples from Hungarian are relevant to the points just raised. Hungarian is strongly left-headed phonologically: words have fixed stress on the first syllable, and at least certain kinds of phrases, such as adjective-noun phrases, are left-headed as well.

(5) NEHÉZ nyelv "difficult language" [no contrast implied on *nehéz* "difficult"]

The similarity of word-stress and phrase-stress patterns appears to provide evidence for Halle & Vergnaud's assumption that "stress" is a unified phenomenon, whether at the word level or higher. At the same time, the stress pattern in (5) may also support Nespor & Vogel's claim about the relation between syntactic and phonological headedness, since Hungarian is generally head-final syntactically. However, at the level of the major IP (e.g. long noun phrase or whole sentence), there are plenty of cases where even Hungarian is right-headed, such as long number phrases, discussed by Varga (1998):

(6) KÉTszáz "two hundred"
kétszáz HÁROM "two hundred three"

There are also, of course, other languages where adjective-noun phrases do not so obviously support the preliminary typological claims. English has both left-headed and right-headed aspects to its word-stress rules, so it is at best unclear what Halle & Vergnaud would predict for adjective-noun phrases like *difficult language.* Nespor & Vogel would apparently wrongly predict stress on the adjective, since adjective-noun phrases are head-final in English. Clearly, much research remains to be done.

3.3. Prosodic typology

3.3.1. Syntax-prosody mapping

The term "prosodic" typology is used here to refer to differences in the way prosodic structure is related to morphosyntactic structure. As noted in §2.2. above, universalist descriptions emphasize the naturalness of the association between IP boundaries and morphosyntactic constituent boundaries. The empirical reality appears much messier. For example, there are many well-known cases in which IP boundaries seem to "come in the wrong place", such as the English children's verse *The House that Jack built:*

(7) This is the dog | that chased the cat | that killed the rat | that ate the malt | that lay in the house that Jack built.

As noted by Chomsky & Halle (1968:371f), the IP boundaries (indicated by |) interrupt the repeatedly right-branching noun phrases. The existence of such syntax-prosody "mismatches" has been the subject of a great deal of theoretical research since the 1970s (e.g. Langendoen 1975, Nespor & Vogel 1986). One contribution to this line of research, that of Steedman (e.g. 1991), argues that the problem is not a matter of syntax-prosody mapping at all, but that the complications reside rather in the mapping from semantics to syntax, and in the nature of syntactic structure itself. Most other researchers, however, continue to treat this as a problem of the "interface" between

syntax and phonology, and there has been a considerable amount of explicitly typological work on the topic within the generative principles-and-parameters approach.

In an influential article, Selkirk (1986) proposed that languages vary systematically in (a) whether they locate IP boundaries relative to the left or right edges of syntactic constituents, and (b) which constituents are relevant for IP boundary placement. In the case of *The House that Jack built*, one could say that an IP boundary is placed at the right edge of every syntactic maximal projection; the choice of right edges and maximal projections is a typological generalization about English, and the claim would be that other languages could make different choices. Selkirk's proposal has spawned much descriptive work on a variety of languages (e.g. Rice 1987 on Slave, Chen 1987 on Xiamen Chinese), and has more recently been recast in terms of Optimality Theory (Selkirk 1995, Truckenbrodt 1999), using the notion of "generalized alignment" constraints (McCarthy & Prince 1993).

Again, note that the distinction between "major IPs" and "minor IPs" (§3.2.4) may help to reconcile this typological work with the universalist assumption of a natural connection between syntactic constituent boundaries and IP boundaries. Minor IPs, which are the focus of much of the work based on Selkirk, may show considerable cross-linguistic difference, while major IPs may more consistently correspond to whole sentences and parenthetical phrases.

3.3.2. Sentence-stress and phrasing

It has been observed by Venditti, Jun, & Beckman (1996) that Japanese and Korean exhibit a prosodic phenomenon that is comparable pragmatically, but not phonologically, to English deaccenting (§3.2.2). Contexts which would lead to deaccenting or other sentence-stress shifts in English often lead to what Venditti et al. term *dephrasing* in Japanese and Korean: two content words which would normally each comprise separate prosodic words or "accentual phrases" are combined into a single phrase. This is seen in the following example from Korean (Venditti et al. 1996: 307):

(7) {kwijəun} {maŋatʃi} "cute colt" [two phrases, pragmatically neutral]
{kwijəun maŋatʃi} "CUTE colt" [one phrase, focus on adjective]

One of the words in the combined phrase corresponds to the deaccented word in English, but it is not possible to analyze the Korean data as involving "deaccenting", because accent is lexically assigned. As Ladd (1996) notes, this correspondence between deaccenting and dephrasing suggests that sentence-stress in English and similar languages is not primarily a matter of the location of pitch accents, but more generally a matter of relative prominence within prosodic structure. This possibility needs to be kept in mind in investigating typology, because it may mean that what we are calling "accentual typology" and "prosodic typology" are simply two different facets of the same basic phenomenon. Once again, this is a fertile field for future research.

4. Intonational Phonetics

This section briefly discusses several dimensions of cross-linguistic variation regarding the phonetic realization of intonational features and categories. These are roughly analogous to segmental phonetic differences like whether languages have vowel reduction in unstressed syllables, or whether they have retroflex stops (or clicks, or front rounded vowels). It is possible that variation on these points is systematic, but little definite is known, and the following typological remarks must be regarded as speculative. A further topic that fits naturally in this section, the distinction between stress-timed and syllable-timed rhythm, is discussed separately in article 126.

4.1. Declination

While declination is apparently a universal tendency, the details certainly vary from language to language. Two possible dimensions of variation should be noted. First, it may be useful to think of the extent to which declination is "phonologized" in any given language. Many languages have rules of *downstep*, which cause a given tone to be realized at a lower pitch than a previous occurrence of the same tone. These rules are well attested in the lexical tone languages of Africa (e.g. Clements 1979) and in Japanese (Kubozono 1990), and are argued by many to apply in the European languages as well (Pierrehumbert 1980, Grice 1995). Second, it seems likely that languages differ in how much the overall amount of declination is to be explained by downstep and similar phonological rules, and how much is truly gradual background declination that is part of the overall phonetic shape of an IP. It has been suggested that some languages, especially languages with three or more lexical level tone phonemes, limit background declination severely (e.g. Yoruba, Connell and Ladd 1990).

4.2. Overall pitch range

The term "overall pitch range" can be defined quantitatively in a number of partially contradictory ways; because of this, and because of the conspicuous differences of pitch range between male and female speakers or between "lively" and "monotonous" speakers, it is difficult to make valid statements about pitch range in cross-language perspective. However, impressionistic evidence suggests that *ceteris paribus* speakers of some languages habitually use a wider or higher range than speakers of other languages. For example, the data in Connell and Ladd's study of Yoruba showed that two of the three male speakers used a pitch range comparable to that of the one female speaker, and very much higher than ranges typically reported for male speakers of European languages. It is tempting to speculate that this may be related to Yoruba's use of lexically specified level tone, but other explanations are at least as plausible. In particular, cultural explanations related to gender differences are almost certainly relevant to many such cases, if not to Yoruba itself: it is known that some cultures place a greater value than others on low male voices or high female voices (e.g. Scherer, London & Wolf 1973 on American vs. German; Bezooijen 1995 on Dutch vs. Japanese). In our present state of knowledge it is impossible to assess the respective roles of phonological and cultural factors in cross-language pitch range differences.

4.3. Dynamic vs. melodic accent

Beckman (1986), basing herself on a careful experimental comparison of Japanese and English, shows that the traditional distinction between dynamic and melodic accent has a considerable basis in acoustic reality. In particular, she shows that there is a useful distinction to be drawn between "stress" and "non-stress" accent: in "non-stress" languages (like Japanese) accent is signalled almost exclusively by pitch, and features related to force of articulation (duration, intensity, spectral differences, etc.) are virtually absent; in "stress" languages (like English), duration, intensity, and spectral differences are important concomitants of the pitch features that signal accent, and may even signal accent on their own (i.e. in the absence of pitch cues) in some contexts. Work on Dutch by Sluijter and van Heuven (1996) appears to confirm this view, though there is considerable disagreement about the details (cf. Beckman and Campbell 1997).

For typological purposes, the only important addendum to Beckman's claim is the one made by Ladd (1996). Ladd points out that Beckman's choice of English and Japanese as examples may give rise to the impression that "stress" languages have only intonationally specified pitch features (cf. §3.1.), while pitch features in "non-stress" languages are lexically specified. The two typological dimensions are actually independent. Both Swedish and Chinese have (in rather different ways) lexically specified pitch features, but both are also widely said to have stress. Other languages

(such as Bengali and Indonesian) certainly have intonationally specified pitch features, and probably have lexical accent (in the sense that some syllable in each word is designated the bearer of any intonational pitch prominence), but do not appear to make any use of stress in Beckman's or Sluijter and van Heuven's sense. Again, this is a rich potential field for typological research.

5. References

Abramson, Arthur (1962). The vowels and tones of standard Thai: Acoustical measurements and experiments. *International Journal of American Linguistics* 28, No. 2 part 3.

Arvaniti, Amalia; Ladd, D. Robert; Mennen, Ineke (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26, 3-25.

Beckman, Mary E. (1986). *Stress and Non-Stress Accent*. (Netherlands Phonetic Archives 7). Dordrecht: Foris Publications.

Beckman, Mary E. & Pierrehumbert, Janet B. (1986) Intonational Structure in English and Japanese. *Phonology Yearbook* 3, 255-310.

Bezooijen, Renée van (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech* 38, 253-265.

Bolinger, Dwight (1972). Accent is predictable (If you're a mind-reader). *Language* 48, 633-644.

Bolinger, D. 1978. Intonation across languages. In J. Greenberg (ed.) *Universals of Human Language, Volume 2: Phonology*. Stanford: Stanford University Press, pp. 471-524.

Bolinger, Dwight (1986). *Intonation and its Parts*. Stanford: Stanford University Press.

Bolinger, Dwight (1989). *Intonation and its Uses*. Stanford: Stanford University Press.

Brown, Gillian; Currie, Karen; Kenworthy, Joanne (1980) *Questions of intonation*. (London: Croom Helm).

Bruce, Gösta (1977). *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.

Campbell, W. N.; Beckman, Mary E. (1997). Stress, prominence, and spectral tilt. In A. Botinis, G. Kouroupetroglou, and G. Carayiannis (eds.), *Intonation: Theory, Models and Applications*. European Speech Communication Association.

Chang, Nien-Chuang (1958). Tones and Intonation in the Chengtu Dialect (Szechuan, China). *Phonetica* 2, 59-84. Reprinted in D. Bolinger (ed.), *Intonation* (Penguin Books, 1972), pp. 391-413.

Chen, Matthew (1987). The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4, 109-149.

Cinque, Guglielmo (1993). A null theory of phrase and compound stress. *Linguistic Inquiry* 24, 239-297.

Chomsky, Noam & Halle, Morris (1968). *The Sound Pattern of English*. New York: Harper and Row.

Clements, G. N. (1979). The description of terrace-level tone languages. *Language* 55, 536-558.

Cohen, A., Collier, René, & 't Hart, Johan (1982). Declination: construct or intrinsic feature of speech pitch? *Phonetica* 39, 254-273.

Connell, Bruce, & Ladd, D. Robert (1990). Aspects of pitch realisation in Yoruba. *Phonology* 7, 1-30.

Cooper, William, & Sorensen, John (1981). *Fundamental Frequency in Sentence Production*. Heidelberg: Springer.

Cruttenden, Alan (1981). Falls and rises: Meanings and universals. *Journal of Linguistics* 17, 77-91.

Cruttenden, Alan (1993). The de-accenting and re-accenting of repeated lexical items. *Lund University Working Papers in Linguistics and Phonetics* 41, 16-19.

Cruttenden, Alan (1994). Rises in English. In J. Windsor Lewis (ed.), *Studies in General and English Phonetics: Essays in Honour of Professor J. D. O'Connor.* (London: Routledge), pp. 155-173.

Fujisaki, Hiroya (1983). Dynamic characteristics of voice fundamental frequency in speech and singing. In Peter F. MacNeilage,(ed.), *The Production of Speech* (New York: Springer), pp. 39-55.

Gårding, Eva (1987). Speech act and tonal pattern in standard Chinese: constancy and variation. *Phonetica* 44, 13-29.

Grice, Martine (1995). Leading tones and downstep in English. *Phonology* 12, 183-233.

Grønnum, Nina (1991). Prosodic parameters in a variety of regional Danish standard languages, with a view towards Swedish and German. *Phonetica* 47, 188-214.

Grønnum, Nina (1995). Superposition and subordination in intonation: A non-linear approach. *Proceedings of the 13th International Congress of Phonetics Sciences, Stockholm*, vol. 2, pp. 124-131.

Guierre, Lionel (1979). *Essai sur l'accentuation en anglais contemporain*. Université de Paris VII.

Gussenhoven, Carlos; Bruce, Gösta (1999). Word prosody and intonation. In H. van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe* (Berlin: Mouton de Gruyter), pp. 233-271.

Halle, Morris, & Vergnaud, Jean-Roger (1987). *An Essay on Stress*. Cambridge MA: MIT Press.

't Hart, J., Collier, René, & Cohen, A. (1990*). A perceptual study of intonation: An experimental-phonetic approach.* Cambridge: Cambridge University Press.

Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Kubozono, Haruo (1989). Syntactic and rhythmic effects on downstep in Japanese. *Phonology* 6, 39-67.

Ladd, D. Robert (1984). Declination: a review and some hypotheses. *Phonology Yearbook* 1, 53-74.

Ladd, D. Robert (1990). Intonation: emotion vs. erammar (Review article on Bolinger 1989). *Language* 66, 806-816.

Ladd, D. Robert (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Ladd, D. R., Faulkner, D. S., Faulkner, H. & Schepman, A. (1999). Constant "segmental anchoring" of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America* 106, 1543-1554.

Langendoen, D. Terence (1975). Finite-state parsing of phrase-structure languages and the status of readjustment rules in the grammar. *Linguistic Inquiry* 6, 533-554.

Liberman, Mark; Pierrehumbert, Janet (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff and R. Oerhle (eds.) *Language Sound Structure* (Cambridge MA: MIT Press), pp. 157-233.

Lieberman, Philip (1967). *Intonation, Perception, and Language*. Cambridge MA: MIT Press.

Lindau, Mona (1986). Testing the Model of Intonation in a Tone Language. *Journal of the Acoustical Society of America* 80: 757-764.

McCarthy, John & Prince, Alan (1993). Generalized alignment. In G. Booij & J. van Marle (eds.), *Yearbook of Morphology 1993*, pp. 79-153.

Möbius, Bernd, Pätzold, Matthias, & Hess, Wolfgang (1993). Analysis and synthesis of German F0 contours by means of Fujisaki's model. *Speech Communication* 13, 53-61.

Nespor, Marina, & Irene Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Odé, Cecilia, & van Heuven, Vincent J. (eds.) (1994). *Experimental Studies of Indonesian Prosody.* University of Leiden.

Ohala, J. J. (1982). Cross-language use of pitch: An ethological view. *Phonetica* 40, 1-18.

Pierrehumbert, Janet (1980). *The Phonology and Phonetics of English Intonation*. PhD Dissertation, Massachusetts Institute of Technology.

Pierrehumbert, Janet & Beckman, Mary (1988). *Japanese Tone Structure*. Cambridge MA: MIT Press.

Pike, Kenneth L. (1945). *The Intonation of American English*. Ann Arbor: Univ. of Michigan Press.

Prieto, Pilar, van Santen, Jan & Hirschberg, Julia (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics* 23, 429-451.

Rice, Keren (1987). On defining the intonational phrase: evidence from Slave. *Phonology Yearbook* 4, 37-60.

Scherer, K. R., London, H., & Wolf, J. J. (1973). The voice of confidence: paralinguistic cues and audience evaluation. *Journal of Research in Personality* 7, 31-44.

Selkirk, Elisabeth O. (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3, 371-405.

Selkirk, Elisabeth O. (1995) Sentence prosody: Intonation, stress, and phrasing. In J. Goldsmith (ed.), *The Handbook of Phonological Theory* (Oxford: Blackwell), pp. 550-569.

Sluijter, Agaath, & van Heuven, Vincent (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100, 2471-2485.

Steedman, Mark (1991). Structure and intonation. *Language* 67, 260-296.

Thorsen, Nina (1980). A study of the perception of sentence intonation: Evidence from Danish. *Journal of the Acoustical Society of America* 67, 1014-1030.

Truckenbrodt, Hubert (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219-255.

Vaissière, Jacqueline (1983). Language-independent prosodic features. In A. Cutler and D. R. Ladd (eds.), *Prosody: Models and Measurements* (Heidelberg: Springer), pp. 53-66.

Vallduví, Enric (1992). *The Informational Component*. New York and London: Garland Press.

Varga, László (1998). Rhythmical variation in Hungarian. *Phonology* 15, 227-266.

Venditti, Jennifer, Jun, Sun-Ah & Beckman, Mary E. (1996). Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English. In James L. Morgan and Katherine Demuth (eds*.), Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Hillsdale NJ: Lawrence Erlbaum), pp. 287-311.

Zubizarreta, Maria Luisa (1998). *Prosody, Focus, and Word Order*. Cambridge MA: MIT Press.