Suprasegmentals

Carlos Gussenhoven, Radboud University Nijmegen, Nijmegen, The Netherlands

Abstract

The prosodic structure of languages is divided into word prosody and sentence prosody. Beyond an organization of vowels and consonants into syllables, languages may have segmental length, tone, and stress, in any combination. The prosodic structure of words can be inferred from their pronunciation by abstracting away from the phonological phrasing structure (the prosodic hierarchy) and the tonal structure of the intonational melody. Boundary tones delimit larger phonological phrase types and some languages have phrase-internal tones (pitch accents). Word prosody distinctions encode lexical and grammatical morphemes, prosodic phrasing imperfectly reflects syntactic structure, and melodic contrasts signal discourse and focus meanings.

Introduction

The prosody of language covers all aspects of speech that are not related directly to the articulation of the vowels and consonants in linguistic expressions, a negative definition that is echoed by the traditional term 'suprasegmentals.' The topic can be approached in positive terms either from the perspective of the phonetic signal or from that of the linguistic structure. The first perspective leads to the analysis of the acoustic signal into four dimensions of variation. These are the spectrum, duration, intensity, and fundamental frequency. It was pointed out by Ilse Lehiste (1970) that the last three dimensions can be observed as properties of the first. We can say how long and how intense the speech signal for some vowel is and what its fundamental frequency is, but we cannot say how intense duration is or how long the fundamental frequency, she argued. In this sense, suprasegmental features are overlaid features, while spectral features are inherent. The second perspective presupposes an understanding of the prosodic structures of languages. It seems fair to say that this understanding has not yet achieved the level of analytic accuracy that has been reached for segmental structures, but considerable progress has been made in the past decades.

Two general points are made before we move on to linguistic structures. One is that the neat distinction between spectral variation and the three suprasegmental dimensions of variation must blur again, because phonological elements are famously promiscuous with respect to the four dimensions of phonetic variation. Thus, linguistic categories that typically are encoded in one of the suprasegmental dimensions of variation frequently also make use of spectral variation, including variation in voice quality. This applies to stress and tone as well as to intonation. The second point is that a great deal of the communicative impact of intonated speech is to be attributed to paralinguistics, the systematic variation that speakers of all languages use to signal universal meanings not encoded in the linguistic structure (Ladd, 2008). Variations in overall pitch range, for example, may signal such general meanings as surprise, authoritativeness, and emphasis (Ohala, 1994). Generally, paralinguistic meanings can be seen as metaphorical interpretations of anatomical and physiological effects on vocal fold vibration (Gussenhoven, 2004, Chapter 4). Ohala's Frequency Code connects the effect of larynx size on frequency of vocal fold vibration with the inferred size of the animal or human producing it, causing high pitch to signal 'small' meanings like submissiveness, vulnerability, and uncertainty and low pitch to signal their opposites. The disentanglement of these paralinguistic meanings – from the meanings encoded in intonational morphemes – will remain an intriguing area of research. A related point is that intonation differs from other components of linguistic structure in that deviations in the timing or pitch range from a canonical pronunciation will trigger paralinguistic interpretations of speaker intentions, while comparable spectral deviations may be interpreted just as indications of articulatory casualness or of temporary physical constraints on speech production.

The approach adopted in the remainder of this article takes the phonological representation as the starting point, selecting those elements in that representation whose realization typically or in large measure involves nonspectral variation, much as has been done in Fox (2000). Since few languages will exemplify all the prosodic phenomena that have been reported, this approach inevitably looks at the languages of the world as a single body of data and will identify aspects of prosodic structure cross-linguistically. One way to divide the topic is to separate word prosody from sentence prosody.

Word Prosody

Broadly, the phonological structure of words can be more than a syllabified string of consonants and vowels in three ways. One is the existence of long segments by the side of short ones, a second is the presence of metrical structure, and a third is the presence of lexical tone. Languages do not require any of these properties. French, for instance, has been claimed to be a language without any word prosodic structure, which is supported by the finding that French listeners perform poorly in short-term memory tasks involving the position of the stress (Peperkamp and Dupoux, 2002).

Quantity

When duration differences are encoded in the phonology, they are referred to as quantity differences or length differences. Vowels as well as consonants may contrast for quantity. Finnish has quantity contrasts for both consonants and vowels, as has Japanese. Hawaiian contrasts long and short vowels only, while Italian contrasts long and short consonants only. Spanish and French have neither long vowels nor geminate consonants. Quantity is described with the help of moras, which make up the syllable rime, the part of the syllable minus the initial (onset) consonant(s), but including the closing (coda) consonants. A Finnish syllable rime containing a short vowel, like the first syllable of *eri* 'different,' has one mora, while a syllable with a long vowel, like the first syllable of *vaara* 'danger,' has two. A long vowel is prosodically identical to a short vowel plus a coda consonant, making *vaara* and *arki* 'workday' prosodically equivalent. Three moras occur in the first syllable of *arkki* 'sheet of paper' or *aarre* 'treasure,' where the first half of the geminate consonants [kk] and [rr] is a coda in the first syllable and the second half an onset in the second. Syllables with four moras, like [aark] in the proper name *Jotaarkka*, are extremely rare (Suomi et al., 2008). Japanese syllables have one, two, and very rarely three moras, as illustrated by, respectively, [o] 'tail,' [on] 'speech sound,' and [aan] in (*Nyuu*) *Jiraando* '(New) Zealand.' Trimoraic rimes in which all three moras are filled by the same segment are extremely rare and have been reported for Estonian and related languages and for Dinka. The most frequent phonological context of geminates is intervocalic. A language with word-initial, word-medial, and word-final geminates has been reported in Kraehenmann (2001), together with a discussion of their moraic representation.

Word Prosodic Metrical Structure

A language has metrical word structure if the syllables in a word differ in prominence or weight from other syllables in the same word. In such cases, the more prominent syllable is stressed and the less prominent ones are unstressed. Stress systems show a number of tendencies. One of these is alternation, the tendency for stressed and unstressed syllables to alternate. Another is culminativity, the restriction to a single most prominent syllable. When there are two or more stressed syllable, the most prominent one is referred to as the primary stress (aka main stress, word stress) and the other or others as secondary stress. A third tendency is for the main stress to occur at or near a word edge, the demarcative tendency. These three aspects are illustrated by Pintupi (Australia), where words always have the primary stress on the first syllable, while in careful pronunciation, secondary stresses occur on odd syllables after the first, as in the 12-syllable ['juma,ɻiŋkun,tama,ɻaca,nampa,luɻa] 'avoid (the camp) of all those mothers-in-law' (stresses inferred, Hansen and Hansen, 1969). To capture this alternation, stress systems are described in terms of feet, disyllabic strong-weak (trochee) or weak-strong (iamb) templates, which are assigned sequentially through the word. Pintupi is analyzed as having trochees assigned from left to right, which accounts for the fact that any odd numbered word-final syllable is not stressed (i.e., the last two syllables of words with an odd number of syllables are unstressed). Culminativity and demarcativeness are described by a higher level prominence assignment to either the left or right. Further complexity may be due to interdependencies between stress and the segmental composition of the syllable. In quantity-sensitive languages, long vowels, or long vowels and closed syllables, do not tolerate being in a weak position, and thus will interrupt any regular count through the word. Hawaiian is a quantity-sensitive trochaic language that assigns feet right-to-left, so that [ko'hola] 'reef' has penultimate stress, but [koho'laa] 'whale' has final stress (Elbert and Pukui, 1979). Another type puts no limit on the distance of the heavy syllable from the word edge (unbounded feet). Classical Arabic has initial stress, as in ['balaħatun] 'date,' but if there is a later heavy syllable, the rightmost one has the stress, as in [manaa'diilu] 'kerchiefs.'

At the right edge, the most frequent location of main stress is penultimate (e.g., Mohawk, Polish, Quechua, Swahili), but final and antepenultimate locations are common, often in the same language. On the left, the most frequent position is initial (e.g., Bengali, Czech, Finnish, Icelandic), but stress on the second syllable is common. Extrametricality is a descriptive device used to explain regular or exceptional stress locations further away from the edge. Exceptional antepenultimate stress in languages with regular penultimate stress can be explained by making the last syllable or the last consonant 'invisible' to foot assignment. The opposite ploy is called catalexis, the postulation of an extra syllable at the word end, which is presumed to be parsed as a weak syllable in a foot. Languages may have lexicalized stress within an otherwise-regular stress system, in which case stress will be lexically contrastive, as in English *'insight* vs *in'cite*. Stress location may be sensitive to morphological structure, as in Spanish *'ablo* vs *a'blo* or English *'overhaul* (noun) vs *over'haul* (verb). It generally is believed that no language *only* has lexicalized stress, such that stress locations do not yield to any generalization. A recent survey is van der Hulst et al. (2010).

From a distributional point of view, the notion stress is reasonably uncontroversial. Hyman (2006) characterized a language as having 'stress' if it occurs on a syllable, as opposed to a mora; if it is culminative, it being the most prominent syllable in the word; and if it is obligatory, meaning that there should be no words without stress, discounting unstressed words that cannot appear in an utterance without stressed words, like the English articles. From the perspective of its phonological and phonetic substance, stress is more elusive. On the assumption that the origin of stress is a rhythmic alternation of hyper- and hypoarticulation, it is not surprising that stressed syllables are typically longer and pronounced with greater articulatory care than unstressed syllables, resulting in more canonical articulations with less articulatory reduction and a sustained amplitude across the spectrum (e.g., Sluijter et al., 1997), which enhances intelligibility (Lu and Cooke, 2009).

In many languages, hyperarticulation tendencies are reflected in the phonology. English is a good example of a language in which the composition of a stressed syllable is phonologically different from an unstressed one. Unstressed syllables usually have a three-way vocalic distinction, with details depending on the dialect (Bolinger, 1986: p. 347 ff), as

illustrated by the vowels in the second syllables of *cranial*, *usual*, and *mineral*. Any other vowel would attract the stress, as in *denial*, *bestowal*, *betrayal*, and so on. Second, the stressed syllable in English acts as the attractor of the intonational pitch accent. Less than 40% of the syllables in English prose texts are stressed, and in a competent reading, only about 65% of these are pitch accented. Stressed syllables of other languages may have similar phonological features distinguishing them from unstressed syllables. In Egyptian Arabic, stressed syllables are always pitch accented, for instance (Hellmuth, 2007). In Finnish, which has a quantity distinction on both consonants and vowels and primary stress on the first syllable, all vowels that appear in stressed syllables also can appear in unstressed syllables. (According to Suomi et al. (2008: p. 77), some of the more complex syllable structures occur only in initial position.) If languages have no segmental distinction between stressed and unstressed syllables, the stressed syllable may be variably pitch accented (Beckman and Edwards, 1994), or else the difference is phonetic. Spanish distinguishes unaccented stressed and unstressed syllables with the help of a phonetic duration difference (Ortega-Llebaria and Prieto, 2010).

A fairly widely honored, but deeply confusing, tradition in English phonology is to collapse unaccented stressed and pitch-accented stressed syllables into a single paradigm of syllables with different degrees of stress. The presence of a pitch accent on a word's stressed syllable does not alter any of its segmental features, but it adds pitch features in and around the accented syllable that make up the intonation contour of the utterance (Bolinger, 1986). The determination of what is a stressed syllable belongs virtually exclusively in the lexicon; a minor role for syntax is found in the obligatory stressed pronunciation of function words before deletion sites, as illustrated in *Sue can* (kən) *come, but I don't know if Gary can* (kæn) (Kaisse, 1985; Selkirk, 1986, 1995). The reasons for the presence of pitch accents on words are more varied. They are morphological, phonological, or syntactic or they have to do with information structure (cf Gussenhoven, 2011). Compound formation may be marked by lack of a pitch accent in the second constituent, as in *COURSE requirements*, while phonological phrasing is responsible for the difference in accentuation in *The gave [the CHInese DISHES]φ* ('They gave the dishes that were of Chinese origin') *They gave [the CHINESE]φ [DISHes]φ* ('They gave dishes to the Chinese') (Hayes, 1989). Syntax is involved in the generalization that predicates may lack a pitch accent when adjacent to an argument. The predicate–object combination in *I used chop-sticks* typically is pronounced with a pitch accent on *CHOPsticks* only, while the noun phrase *USED CHOPsticks* will have two pitch accents (Schmerling, 1976; Selkirk, 1986; Birch and Clifton, 1995). A condition on the deletion of the accent on the verb is that the proposition should be eventive (Gussenhoven, 1983; cf the distinction between stage-level and individual-level predicates in Kratzer, 1995). For instance, the verb in *Too many cooks spoil the broth* has an obligatory pitch accent (in addition to the pitch accents on *Too*, *cooks*, and *broth*), because the sentence is a 'contingency' sentence ('If there are …' etc.). The absence of that pitch accent would invite the interpretation that the number of cooks that spoil the broth is in fact large. Information structure is involved in the difference between *Is that an ENTrance requirement or a COURSE requirement?* In the literature that collapses the presence of a pitch accent with word-stressed syllables, the compound *COURSE requirements* would be said to have secondary stress on the penultimate syllable of *requirements* and primary stress on *course*, while in the case of the contrastive focus on *course* in *COURSE requirement*, the word *requirement* would be said to be deaccented. The prosodic structures of these two expressions are the same, however.

Lexical Tone

A frequently applied morphosyntactic criterion for lexical tone, and hence of the notion *tone language*, is whether tone(s) enter into the specification of words or morphemes. If they do, the language is a *tone language*. Examples are found in Asia (Sino-Tibetan languages, Japanese), Africa (Bantu languages, Khoi-San languages, Nilo-Saharan languages), and many languages in the Americas (Yip, 2002). Mandarin (Sino-Tibetan) has five words that consist of the syllable [ma]. Four of these have a lexical tone: [ma˥] 'mother,' [ma˧˥] 'hemp,' [ma˨˩] 'horse,' [ma˥˩] 'scold,' where the Chao tone letters give an iconic indication of the pitch contour to the left of a vertical bar, which indicates the pitch span (Chao, 1930). One [ma] is toneless, a question particle ('neutral tone'). Notation of tones in tone languages spoken in Africa and, to a lesser extent, in the Americas use accents over vowels (high level [á], low level [à], mid level [ā] falling [â], rising [ǎ], etc.), whereby long vowels are given as [aa], as in Mamvu [máàkà] 'type of seasoning' vs [màákà] 'cat.' Frequently, tone is involved in the morphology, as in Makonde [nìndì tálèékà] 'I cooked' vs [nìndà tàlèékà] 'I will cook' and Mamvu [áfù] 'man,' [àfú] 'men' (examples from Kaji, 1999). Estimates of the proportion of tone languages run from 50% (Maddieson, 1978) to 70% (Yip, 2002).

A subclass of tone languages frequently is singled out as being pitch-accent languages. Hyman (2009) argued that languages that have been identified as such differ from other tone languages in the distribution of the tone contrast over the word. Tone contrasts typically are not located on every syllable or mora of the word. Sino-Tibetan languages may have very 'dense' distributions, with only a minority of words, like the interrogative particle [ma] of Mandarin, having no tone of their own ('neutral tone'). Japanese is an example of a language with a 'sparse' distribution of tone. Only about half of the Japanese words, the accented words, have a lexical pitch accent (HL), the other half being lexically toneless, or unaccented. Since the pitch accent can be lexically specified once per word for any syllable, the number of prosodic patterns for words of n syllables is $n - 1$. A triplet of disyllables is [ha˥si] 'chopsticks,' [hasi˥] 'bridge,' and [hasi] 'end,' where the right hook indicates a pitch accent on the preceding syllable.

The study of tonal structures in African languages in the second half of the last century led to the development of autosegmental representations (Goldsmith, 1976), whereby the phonological specification of words consists of at least two parallel linear tiers, in which the elements (the 'segments') are synchronized by means of associations. In

the case of the synchronization of tones on the one hand and vowels and consonants on the other, the associations are not direct, but rather they are mediated by the prosodic structure. Some language use the syllable as the tone-bearing unit (TBU), others the mora, whereby the tone-bearing status of the mora is usually conditional on the presence of a vowel or sonorant consonant in it, thus excluding obstruent consonants from being tone bearers. Although the Japanese pitch accent is assigned to a syllable, the tones associate with moras. That is, even though the number of moras in [hai] is the same as that in [hasi], there could be only two words with the monosyllabic structure, [hai] and [hai¬], both of which happen to exist, 'lungs' and 'yes,' respectively. The associations of the lexical pitch accents with the previous five words are shown in (1) for the disyllables and (2) for the monosyllables. Structurally, there is no (2b).

(1) a. h a s i — H L — 'chopsticks'
b. h a s i — H L — 'bridge'
c. h a s i — 'edge'

(2) a. h a i — H L — 'yes'
b. -
c. h a i — 'lungs'

Autosegmental representations explain, as in (1a) and (2a), how the same tone string, here HL, can have different syllabic pitch effects. The first syllable in (1a) is high, the second is low, and the syllable in (2a) has a fall. For this reason, linguists are careful to distinguish phonological tones from their phonetic realization.

Tone languages spoken in Africa have varying degrees of density. Somali is comparable to Japanese in that it has maximally one lexical H-tone (or 'pitch accent') per word, but it is assigned to the mora, which means that a monosyllable like [ei] can have H either on the first mora ('dog') or on the second ('dogs'). Unaccented moras have L-tone. The location and presence of the H pitch accent in Somali largely signal morphological and syntactic, as opposed to lexical information. In addition to marking number in some nouns by its location, the H pitch accent is deleted in WH-phrases, for instance, as in [ei geé] 'which dog'? (Hyman, 1981). Other languages have a number of word melodies, like H, HL, LHL, that are specified along with the vowels and consonants for each word. The way tone strings distribute themselves over the TBUs of the word formed a significant motivation behind the development of autosegmental phonology, whose essential claim is that more than one segment on one tier may correspond to a single unit on the other. This is illustrated for the Mende words in (5) (from Leben, 1978). All three words happen to have the LH word melody, but because of the different numbers of syllables, their syllabic tone patterns come out as 'rise' for (3a), 'low-high' for (3b), and 'low-high-high' for (3c). This is achieved by assuming that the first tone of the melody associates with the first syllable and subsequent tones with subsequent syllables, if there are any, that leftover tones are associated with the final syllable and leftover syllables are associated with the last tone. (Not all the words of the language follow precisely this pattern of association.)

(3) a. mba — L H — 'rice'
b. navo — L H — 'money'
c. ndavula — L H — 'sling'

Tone grammars can be quite complex. There is a general tendency for tones to be pronounced later than the syllables with which they are lexically associated, which leads to a wide variety of rules for spreading and moving tones rightward. Also, in the way they regulate the associations of tones, grammars may effectively ignore one of the tones, leaving empty TBUs to be filled up with a 'default tone' (Hyman, 2001). Both features are illustrated by Nochixtlán Mixtec, where a final H-tone in morphemes associates with the rightmost mora of the next word. At the same time, all moras must be provided with tones. In (4a), the tone of [kasi^{H}] 'eat' associates with the last syllable of [lana^{MH}]. The preceding H of the verb prefix [nā^{H}] will fill up the first syllable of [lana^{MH}], while the initial three empty syllables are filled with default M (not shown). The final MH of [lana] are deleted. The resulting pronunciation is (4b) (McKendry, 2013).

(4) a. nã kasi lana — H H MH — SUBJUNCTIVE eat child
b. nã̄ kāsī láná — 'May the child eat'

Many two-tone systems in Bantu are underlyingly H vs Ø, as opposed to H vs L, whereby Ø is filled in with L on the surface. The motivation for such privative analyses is provided by the presence of rules that refer to H and the absence of any rules that refer to L (Hyman, 2001). For instance, no morpheme could consist of, contain or assign L, while for H at least one of these operations exists. Whereas H is typically subject to distributional restrictions, for instance, in never occurring in nonperipheral syllables in the underlying word or in being banned from final syllables in the phrase on the surface, the distribution of L can be expressed only in terms of where H does *not* appear. Languages may have the same tone as a default tone and an underlying tone. In Nochixtlán Mixtex, an L-tone between H and M is raised to M, causing HLM to be HMM on the surface. If the third position is empty, however, a preceding HL sequence remains intact, even though an M will appear as a default tone on the surface, giving HLM (McKendry, 2013). If there is a default tone, it is most likely to be L in a two-tone system and M in a three-tone system, but these patterns are not exclusive.

Sentence Prosody

The prosody of the sentence is determined by its prosodic phrasing and its tonal structure. The prosodic phrasing consists of a hierarchically structured set of phonological constituents (Selkirk, 2011; Nespor and Vogel, 1986; Hayes, 1995). The tonal structure has different sources. Boundary tones come with one or more of the phonological constituents, intonational pitch accents, and any lexical tones that are introduced by the words.

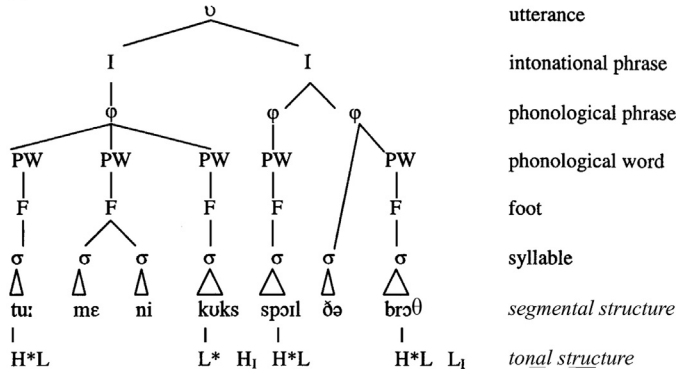
The Prosodic Hierarchy

In (4), an example of a prosodic representation is given of an English sentence, together with frequently used symbols. The boundary tones are subscripted for the phonological constituent they come with, in this case the IP.

H-tone, pronounced after the pitch accent on *cooks*. Boundaries of this type often are felt to correspond with a comma in writing. The utterance, finally, roughly corresponds with the intuitive notion of spoken sentence. There are three phonological phrases (*Too many cooks*, *spoil*, and *the broth*), two intonational phrases (*Too many cooks* and *spoil the broth*), and one utterance (*Too many cooks spoil the broth*).

The higher ranking phonological constituents are less predictable from other aspects of the linguistic structure than the lower ranking ones. The most important factor for the phonological phrase and beyond is the syntactic structure, but great length of constituents may lead to the insertion of constituent breaks. For instance, while *Swans like to swim in the river* may well be one IP, a replacement of *Swans* by *Hippopotamuses* is likely to cause a separate IP to arise for the subject. Conversely, a short length may lead either to restructuring, the merging of constituents that would be expected to be separate

(5)



The entire structure above the syllables is assumed to determine how these syllables are rhythmically organized, just as the structure above the foot does this for feet, and so on. The constituent that dominates the foot is the phonological word. In (4), each foot happens to be dominated by a phonological word. A diagnostic for the phonological word in Germanic languages is syllabification: within this constituent, consonants syllabify as onsets with vowels to their right. The phonological word may be smaller or larger than the morphological word. In Germanic languages, each lexical constituent is a separate syllabification domain (e.g., English *cat's eyes* syllabify as [kæts.aɪz], not as *[kæt.saɪz], cf the single phonological word *capsize*). Clitics are morphemes that are included in a following or preceding phonological word (e.g., English *What is he* [wɒt.si] *saying?*). The phonological phrase may determine the distribution of accents, favoring first and last accents. Although, *Too many*! will have two accents, *Too many* COOKS is likely to lose the pitch accent on *many*.

The term 'stress shift' is applied to cases in which such rhythmic deaccentuation leads to the sole presence of a pitch accent on a foot with secondary stress, as in JAPANESE, but JAPANESE FURNITURE (Gussenhoven, 2011). The intonational phrase typically comes with *boundary tones* at its edges. The intonational phrase *Too many cooks* in (4) ends with an

on the basis of the syntax, or cliticization, the inclusion of a short syntactic constituent into a lower ranked constituent than it would have formed by itself if syntax were the only determinant. An example of restructuring occurs in verb–adverb combinations in American English (Nespor and Vogel, 1986, p. 178). If the adverb is a single phonological word, it is likely to merge with the verb into a single phonological phrase. Thus, in *Rabbits reproduce quickly*, the words *reproduce* and *quickly* will combine in a single phonological phrase, while in *Rabbits reproduce quickly and quietly* the verb will retain a phonological phrase to itself. In some varieties of English, the difference is shown by the pronunciation of *reproduce*. In the first example, a pitch accent will occur on *re-*, but in the second, there will (also) be one on *-duce*. In verb–object combinations, a similar situation exists. Thus, *spoil the broth* may form a single phonological phrase, as shown by varieties that allow stress shift in verb–object combinations where the object is a single phonological word. The pronunciation of the verb in MAINTAIN ORDER will then contrast with that in ORDER TO MAINTAIN. In *We need to maintain order and discipline*, however, the pronunciation *MAINTAIN is ungrammatical in those same varieties. In English, cliticization applies to pronouns and auxiliaries, for instance, as illustrated by the pronunciation of *What is he* in the previous paragraph. There is no consensus on how many

prosodic constituents there are across the languages of the world and terminology will vary. Japanese has an *accentual phrase*, which seems to lie halfway between PW and the phonological phrase, also referred to as a *minor phrase*, while a *clitic group* has been claimed for Italian and English, for instance.

Preboundary lengthening refers to the phenomenon that the last syllable of a higher phonological constituent is longer than a medial syllable, an effect that is larger as the constituent rank is higher. At the level of the phonological word in English, it is responsible for the difference in pronunciation between *sin tax* and *syntax*, where [sɪn] in the first structure is longer than in the second. And at the higher end, the insertion of an intonational phrase boundary after *go* in *Let's go and have dinner* will considerably lengthen *go* (e.g., Beckman and Edwards, 1990; Wightman et al., 1992; Gussenhoven and Rietveld, 1992; Byrd et al., 2006).

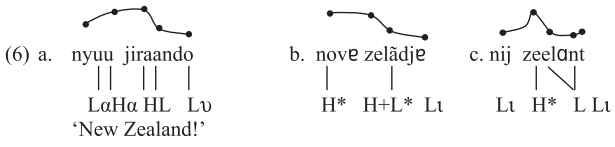
Initial strengthening is the more careful pronunciation of initial segments of a phonological constituent. For instance, the utterance-initial [t] in *Tie it up*! is likely to have a wider area of contact, a longer period of aspiration (a longer, positive Voice Onset Time) than the word initial one in *Stronger ties*, while the same will be true for the latter [t] relative to the foot-initial one in *Valentine* (Cho and Keating, 2001; Keating et al., 2003; Cho et al., 2007).

Finally, the constituents in the prosodic hierarchy often define the context of segmental restrictions. For instance, syllable final plosives and affricates are voiceless in Turkish; syllable final obstruents are voiceless in Dutch; within Italian prosodic words, intervocalic [z] occurs to the exclusion of [s]; and in Kimatuumbi, all vowels are shortened except for any long vowel that ends up in the last word in the phonological phrase (Nespor and Vogel, 1986).

Tone Structure

The tones in a sentence have three sources. First, there are the tones that come with the morphemes in the structure (i.e., the lexical tones). Second, phonological constituents like the intonational phrase and the phonological phrase may begin or end with boundary tones, and third, intonational pitch accents may occur on certain words. Boundary tones will signal the prosodic phrasing of utterances, and since this largely reflects the morphosyntactic structure, will play a crucial role in parsing speech (Cutler et al., 1997; Carlson, 2009; Pannekamp et al., 2005). In addition, there may be a choice of boundary tones, in which case, they signal discourse meanings like 'statement,' 'finality,' 'question,' and so on. Japanese sentences, for instance, have tones from the first two sources (Pierrehumbert and Beckman, 1988). In addition to the lexical pitch accents, the accentual phrase invariably begins with an LH melody, notated LαHα, and the utterance is ended by either an L, signaling declarative meaning, or H, for interrogative meaning. Most European languages derive their tones from the second and third sources, in each case providing a large number of choices, giving a wide variety of discourse meanings (e.g., Pierrehumbert, 1980; Grice, et al. 2005; Frota, 2000; Ladd, 2008; Gussenhoven, 2005; Chen et al., 2007). Swedish and Norwegian have tones from all three sources (Kristoffersen, 2000; Riad, 2013), as do Northern Bizkaian Basque (Hualde et al., 1994), Belgrade Serbian (Smiljanić, 2002), and the dialects in Germany, Belgium, and the Netherlands known as Central Franconian (*Mittelfränkisch*) and Limburgish (e.g., Gussenhoven and Peters, 2004) (cf van der Hulst, 1999). There is considerable variation in the proportions of intonational and lexical tones. For Standard Chinese, the existence of an interrogative utterance boundary tone has been claimed (Duanmu, 2007), where all other tones are lexical. At the other extreme, there is the existence of a pitch-accent-marked focus governing words like [o] 'even' in Bengali, which otherwise only has intonational tones (Lahiri and Fitzpatrick-Cole, 1999). When all the tones have been sequenced into a string, there may be rule-governed assimilations or deletions of tones, after which the surface representation is converted to tonal targets (Pierrehumbert, 1980). Interpolations may be linear or may have somewhat-convex or convex shapes (Niebuhr, 2007; Barnes et al., 2013).

In (5a), the tones for the Japanese word for *New Zealand* are illustrated. The word has a lexical pitch accent (LH) on the penultimate syllable; the declarative tone is notated Lʊ. Together with the initial boundary tones of the accentual phrase, these tones will associate with the available moras in the word, and the target of each tone is indicated by a bullet. (Many speakers avoid rising pitch on a single syllable, as in *nyuu* in (5), replacing it with high pitch.) In (5b), the Portuguese equivalent is shown, which will have a prenuclear H* pitch accent on the stressed syllable [no] and a nuclear H + L* pitch accent in the stressed syllable [lã], where the leading H-tone is pronounced on the preceding syllable, followed by a boundary L-tone of the intonational phrase (Lɪ). In (5c), the Venlo Limburgish word is shown, which illustrates a case of tone assimilation. The language has a lexical H-tone (Accent 2), which assimilates to L if it occurs on the last mora of the intonational phrase and is immediately followed by Lɪ. Tone bearers are the sonorant moras in accented syllables ([zee] in (5c)) and IP-final syllables ([lɑnt] in (5c)), which require a tonal association. This is the reason why lexical L spreads to the available second mora in [zee], which has been left unused by the H*, which associates to the first mora of [zee]. The phonetic reflex of the combination of lexical L and the final boundary Lɪ is a leveling off or slight reversal of the falling pitch movement, as shown in (5c), which contrasts with the straight fall used for Accent 1.



Thus, just as lexical tones are subject to a tone grammar, so too are intonational tones. The distinction between lexical and postlexical phonology (*see* Phonology) is relevant in that no lexical tone rules can affect intonational tones, but as example (6c) shows, postlexical rules can affect both lexical and intonational tones. As a further example, Japanese has rules that govern the priority of tonal associations in cases in which there are more tones than moras, such as when a [LαHα HL Lʊ] tone string (see (6a)) needs to be pronounced on a monomoraic accented word like [kaˈ] 'passable school

grade' or a bimoraic one like [haiˈ] 'yes' (Gussenhoven, 2004, Chapter 10).

The Expression of Focus

Languages frequently exploit prosodic structure for the expression of *focus*, the marking of information status in sentences (REF Information structure/focus; *see* Information Structure in Linguistics). Frequently, this occurs through the compression of the pitch range over words after the focus constituent. In European languages, this is achieved by the deletion of pitch accents after the focus, with conditions for deletion varying from language to language (Vallduví, 1992; Swerts et al., 2002; Ladd, 2008). In other languages, the words after the focus constituent have reduced pitch range, but the tonal structure remains intact, meaning that lexical distinctions in Chinese are retrievable postfocally. It has been suggested that postfocal compression is an areal feature of Europe and North Asia (Xu et al., 2012). Different meanings or categories of information structure are discussed in Lee et al. (2007). For many languages, an understanding of the prosodic expression of focus structure is impossible without considering the syntactic constraints on focus, because these may implicate the distribution or identity of boundary tones or lexical tones (cf Aboh et al., 2007; Fiedler and Schwarz, 2010). Still, in many languages, information structure may show no prosodic effect of prosody at all (e.g., Zerbian, 2007).

See also: Linguistics: Overview; Phonology; Speech Perception; Speech Production, Psychology of; Speech Recognition and Production by Machines.

Bibliography

Aboh, E.O., Hartmann, K., Hartmann, M., 2007. Focus Strategies in African Languages: The Interaction between Focus and Grammar in Niger-Congo and Afro-Asiatic. Walter de Gruyter, Berlin.

Barnes, T.J., Veilleux, N., Brugos, A., Shattuck-Hufnagel, S., 2013. Tonal center of gravity: a global approach to tonal implementation in a level-based intonational phonology. Laboratory Phonology 3, 337–383.

Beckman, M.E., Edwards, J., 1994. Articulatory evidence for differentiating stress categories. In: Keating, P.A. (Ed.), Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III. Cambridge University Press, pp. 7–33.

Birch, S., Clifton, C.J., 1995. Focus, accent, and argument structure: effects on language comprehension. Language and Speech 38, 365–391.

Bolinger, D., 1986. Intonation and Its Parts: Melody in Spoken Language. Stanford University Press, Stanford, CA.

Byrd, D., Krivokapic, J., Lee, S., 2006. How far, how long: on the temporal scope of prosodic boundary effects. Journal of the Acoustical Society of America 120, 1589–1599.

Carlson, K., 2009. How prosody influences sentence comprehension. Language and Linguistic Compass 3, 1188–1200.

Chao, Y.-R., 1930. ə sistim əv "toun-letəz". Le Maître Phonétique 45, 24–27.

Chen, A., den Os, E., de Ruiter, J.P., 2007. Pitch accent type matters for online processing of intonation status: evidence form natural and synthetic speech. The Linguistics Review 24, 317–344.

Cho, T., Keating, P.A., 2001. Articulatory and acoustic studies of domain-initial strengthening in Korean. Journal of Phonetics 29, 155–190.

Cho, T., McQueen, J.M., Cox, E.A., 2007. Prosodically driven phonetic detail in speech processing: the case of domain-initial strengthening in English. Journal of Phonetics 35, 210–243.

Cutler, A., Dahan, D., van Donselaar, W., 1997. Prosody in the comprehension of spoken language: a literature review. Language and Speech 40, 121–201.

Duanmu, S., 2007. The Phonology of Standard Chinese. Oxford University Press, Oxford.

Elbert, S.H., Pukui, M.K., 1979. Hawaiian Grammar. University Press of Hawaii, Honolulu.

Fiedler, I., Schwarz, A., 2010. The Expression of Information Structure: A Documentation of Its Diversity across Africa. Benjamins, Amsterdam.

Fox, A., 2000. Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals. Oxford University Press, Oxford.

Frota, S., 2000. Prosody and Focus in European Portuguese: Phonological Phrasing and Intonation. Garland Press, New York.

Goldsmith, J., 1976. Autosegmental Phonology. MIT, Doctoral dissertation.

Grice, M., Baumann, S., Benzmüller, R., 2005. German intonation in autosegmental-metrical phonology. In: Jun, S.-A. (Ed.), Prosodic Typology: The Phonology of Intonation and Phrasing. Oxford University Press, Oxford, pp. 55–83.

Gussenhoven, C., 1983. Focus, mode and the nucleus. Journal of Linguistics 19, 377–417.

Gussenhoven, C., 2004. The Phonology of Tone and Intonation. Cambridge University Press, Cambridge.

Gussenhoven, C., 2005. The Phonology of Tone and Intonation. Cambridge University Press, Cambridge.

Gussenhoven, C., 2011. Sentential prominence in English. In: van Oostendorp, M., Ewen, C.J., Hume, E., Rice, K. (Eds.), The Blackwell Companion to Phonology. Wiley- Blackwell, Malden, MA and Oxford, pp. 2780–2806.

Gussenhoven, C., Peters, J., 2004. A tonal analysis of Cologne Schärfung. Phonology 25, 251–285.

Gussenhoven, C., Rietveld, A.C.M., 1992. Intonation contours, prosodic structure and preboundary lengthening. Journal of Phonetics 20, 293–303.

Hansen, K.C., Hansen, L.E., 1969. Pintupi phonology. Oceanic Linguistics 8, 153–170.

Hayes, B., 1989. The prosodic hierarchy in meter. In: Kiparsky, P., Youmas, G. (Eds.), Rhythm and Meter. Academic Press, Orlando, FL, pp. 201–260.

Hayes, B., 1995. Metrical Stress Theory. University of Chicago Press, Chicago.

Hellmuth, S., 2007. The relationship between prosodic structure and pitch accent distribution: evidence from Egyptian Arabic. The Linguistic Review 24, 291–316.

Hualde, J.M., Elordieta, G., Elordieta, A., 1994. The Basque Dialect of Lekeitio. Universidad del País Vasco, Bilbao.

van der Hulst, H. (Ed.), 1999. Word Prosodic Systems in the Languages of Europe. Mouton de Gruyter, Berlin.

van der Hulst, H., Goedemans, R., van Zanten, E. (Eds.), 2010. A Survey of Accentual Systems in the Languages of the World. Mouton de Gruyter, Berlin.

Hyman, L.M., 1981. Tonal accent in Somali. Studies in African Linguistics 12, 169–203.

Hyman, L.M., 2001. Privative tone in Bantu. In: Kaji, S. (Ed.), Proceedings of the Symposium on Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics. Tokyo University of Foreign Studies, Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA), pp. 237–257.

Hyman, L.M., 2006. Word-prosodic typology. Phonology 23, 225–257.

Hyman, L.M., 2009. How (not) to do phonological typology: the case of pitch-accent. Language Sciences 31, 213–238.

Kaisse, E., 1985. Connected Speech: The Interaction of Syntax and Phonology. Academic Press, Orlando, FL.

Kaji, S. (Ed.), 1999. Proceedings of the Symposium on Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics. Tokyo University of Foreign Studies: Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA).

Keating, P., Cho, T., Fougeron, C., Hsu, C., 2003. Domain-initial articulatory strengthening in four languages. In: Local, J., Ogden, R., Temple, R. (Eds.), Papers in Laboratory Phonology, vol. 6. Cambridge University Press, Cambridge, pp. 143–161.

Kraehenmann, A., 2001. Swiss German stops: geminates all over the word. Phonology 18, 109–145.

Kratzer, A., 1995. Stage-level/individual-level predicates. In: Carlson, G.N., Pelletier, F.J. (Eds.), The Generic Book. University of Chicago Press, Chicago, pp. 125–175.

Kristoffersen, G., 2000. The Phonology of Norwegian. Oxford University Press, Oxford.

Ladd, D.R., 2008. Intonational Phonology, second ed. Cambridge University Press, Cambridge, UK.

Lahiri, A., Fitzpatrick-Cole, J., 1999. Emphatic clitics in Bengali. In: René Kager, R., Zonneveld, W. (Eds.), Phrasal Phonology. Foris, Dordrecht, pp. 119–144.

Leben, W.R., 1978. The representation of tone. In: Fromkin, V. (Ed.), Tone: A Linguistic Survey. Academic Press, pp. 177–219.

Lee, C., Gordon, M., Büring, M., 2007. Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation. Springer, Dordrecht, The Netherlands.
Lehiste, I., 1970. Suprasegmentals. MIT Press, Cambridge, MA.
Lu, Y., Cooke, M., 2009. The contribution of changes in F0 and spectral tilt to increased intelligibility of speech produced in noise. Speech Communication 51, 1253–1262.
Maddieson, I., 1978. Universals of tone. In: Greenberg, J. (Ed.), Universals of Human Language, Phonology, vol. 2. Stanford University Press, Stanford, CA, pp. 335–365.
McKendry, I., 2013. Tonal Association, Prominence and Prosodic Structure in Southeastern Nochixtlán Mixtec. Ph.D. dissertation, University of Edinburgh.
Nespor, M., Vogel, I., 1986. Prosodic Phonology. Foris, Dordrecht, The Netherlands.
Niebuhr, O., 2007. The signalling of German rising-falling intonation categories – the interplay; of synchronization, shape, and height. Phonetica 64, 174–193.
Ohala, J.J., 1994. The frequency code underlies the sound-symbolic use of voice pitch. In: Hinton, L., Nichols, J., Ohala, J.J. (Eds.), Sound Symbolism. Cambridge University Press, Cambridge, UK, pp. 325–347.
Ortega-Llebaria, M., Prieto, P., 2010. Acoustic correlates of stress in Central Catalan and Castilian Spanish. Language and Speech 54, 73–97.
Pannekamp, A., Toepel, U., Alter, K., Hahne, A., Friederici, A.D., 2005. Prosody-driven sentence processing: an event-related brain potential study. Journal of Cognitive Neuroscience 17, 407–421.
Peperkamp, S., Dupoux, E., 2002. A typological study of stress 'deafness'. In: Gussenhoven, C., Warner, N. (Eds.), Laboratory Phonology, vol. 7. Mouton de Gruyter, Berlin, pp. 203–240.
Pierrehumbert, J.B., 1980. The Phonology and Phonetics of English Intonation. PhD dissertation. MIT.
Pierrehumbert, J., Beckman, M.E., 1988. Japanese Tone Structure. MIT Press, Cambridge, MA.
Riad, T., 2013. The Phonology of Swedish. Oxford University Press, Oxford.
Rietveld, T., Gussenhoven, C., 1990. Preboundary lengthening and intonational and prosodic constituents. Journal of the Acoustical Society of America 88 (S1), S130.
Schmerling, S.F., 1976. Aspects of English Sentence Stress. University of Texas Press, Austin.
Selkirk, E.O., 1986. Phonology and Syntax: The Relation between Sound and Structure. MIT Press, Cambridge, MA.
Selkirk, E., 1995. Sentence prosody: intonation, stress and phrasing. In: Goldsmith, J. (Ed.), The Handbook of Phonological Theory. Blackwell, Cambridge, MA, pp. 550–569. ISBN: 9780631180623.
Selkirk, E., 2011. The syntax-phonology interface. In: Goldsmith, J., Riggle, J., Yu, A. (Eds.), The Handbook of Phonological Theory, second ed. Blackwell, Oxford, pp. 535–549.
Sluijter, A.M.C., van Heuven, V.J., Pacilly, J.J.A., 1997. Spectral balance as a cue in the perception of linguistic stress. Journal of the Acoustical Society of America 101, 503–513.
Smiljanić, R., 2002. Lexical, Pragmatic and Positional Effects on Prosody in Two Dialects of Croatian and Serbian: an Acoustic Study (Ph.D. dissertation), University of Illinois at Urbana-Champaign.
Suomi, K., Toivanen, J., Ylitalo, R., 2008. Finnish Sound Structure: Phonetics, Phonology, Phonotactics and Prosody. University of Oulu.
Swerts, M., Krahmer, E., Avesani, C., 2002. Prosodic marking of information status in Dutch and Italian: a comparative analysis. Journal of Phonetics 30, 629–654.
Vallduví, E., 1992. The Informational Component. Garland, New York.
Wightman, C., Shattuck-Hufnagel, S., Ostendorf, S., Price, P., 1992. Segmental durations in the vicinity of prosodic boundaries. Journal of the Acoustical Society of America 91, 1707–1717.
Xu, Y., Chen, S.-W., Wang, B., 2012. Prosodic focus with and without post-focus compression (PFC): a typological divide within the same language family? The Linguistic Review 29, 131–147.
Yip, M., 2002. Tone. Cambridge University Press, Cambridge.
Zerbian, S., 2007. Investigating prosodic focus marking in Northern Sotho. In: Aboh, E.O., Hartmann, K., Hartmann, M. (Eds.), Focus Strategies in African Languages: The Interaction between Focus and Grammar in Niger-Congo and Afro-Asiatic. Walter de Gruyter, Berlin, pp. 55–79.