SUPRASEGMENTALS: STRESS, LENGTH, TONE, AND INTONATION.

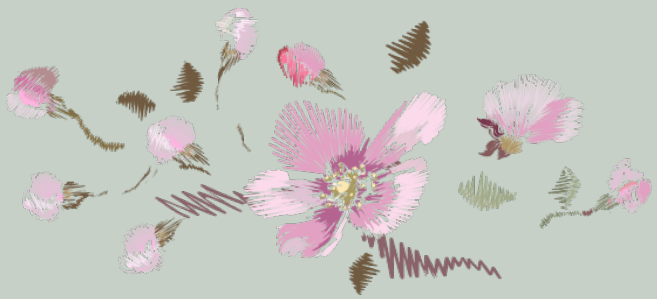
WARQAA MALALLAH HUSSEIN

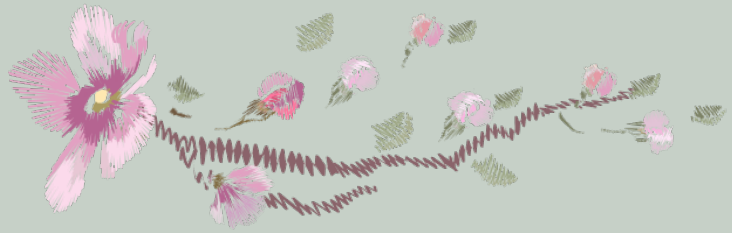
Suprasegmental Features

Spoken language rarely consists of isolated speech sounds (segments). Instead, vowels and consonants are combined to form larger units such as syllables, phrases, and sentences. These larger units can also be described in terms of their articulatory and acoustic properties.

The examination of aspects of speech that extend beyond individual sounds is known as the study of **suprasegmental** or **prosody**.

The principal suprasegmental features are **stress**, **length**, **tone**, and **intonation**.

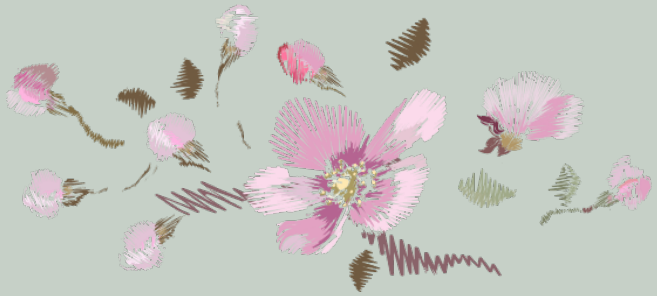
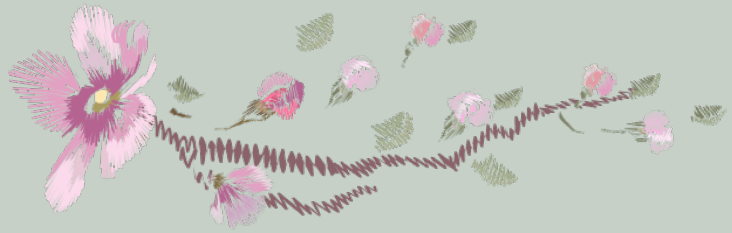




Syllables

A **syllable** is a small unit of speech. Syllables are made up of segments. Syllables often consist of a combination of vowels and consonants, they can also consist of individual vowels or consonants. The syllable is often represented by the Greek symbol **σ**.

For descriptive purposes, a syllable can be divided into **onset** & **rhyme**. The onset is the consonants preceding the vowel. The rhyme of the syllable can be further divided into **nucleus**, which is the vocalic part (vowel), and the **coda** which is the consonants following the vowel.

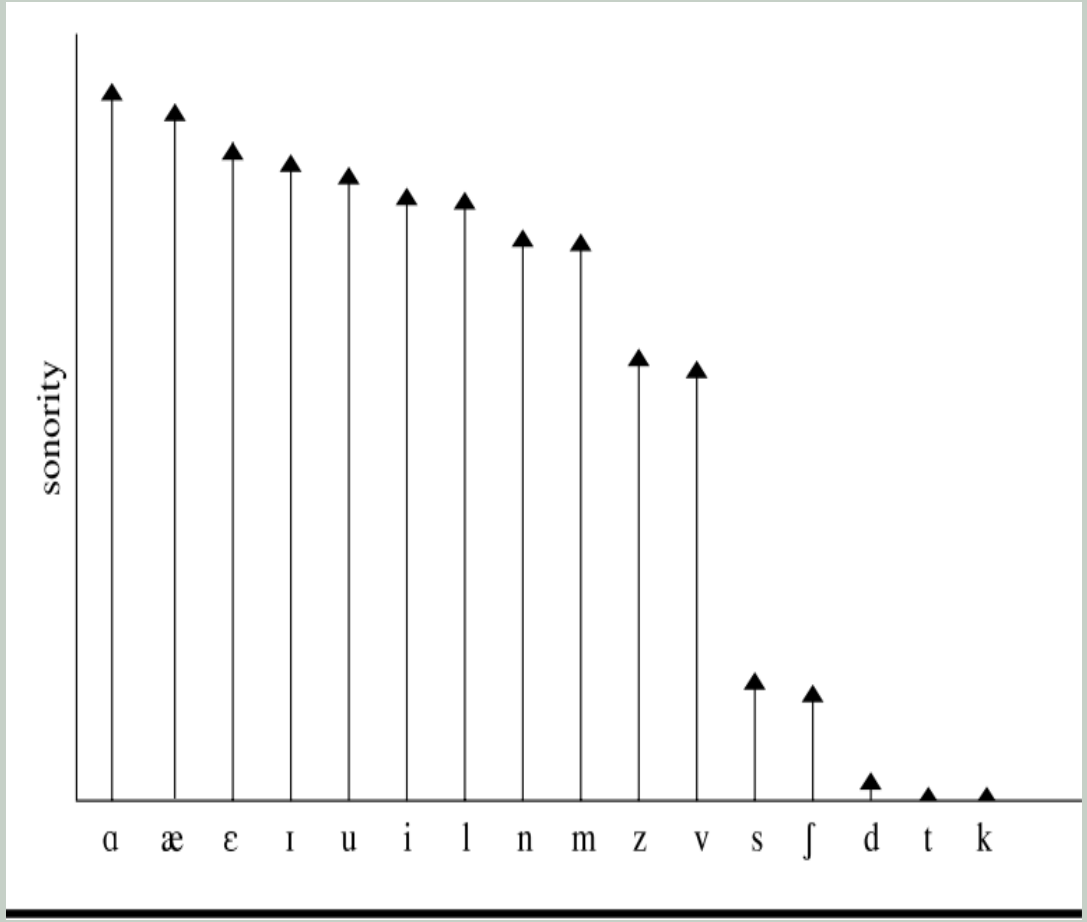

Sometimes it is difficult to say whether a consonant is the coda of one syllable or the onset of another. For example; the word *happy*. Some people will say it is [hæ.pi]; others regard it as [hæp.i]. Another solution is to consider the [p] as belonging to both syllables, they call it **ambisyllabic**. The result would be to transcribe the word as [hæpi] with no syllable division.

Phoneticians disagree on the correct solution to this problem. However, there is a strong tendency to maximize onsets rather than codas and always avoid using ambisyllabic segments. This is known as **the maximal onset principle**. This principle states that where two syllables are to be divided any consonants between them should be attached to the right-hand side syllable not the left, as far as possible, within the restrictions governing syllable onsets and codas.

Phonetic definitions of the syllable have been proposed in both **articulatory** and **auditory** terms.

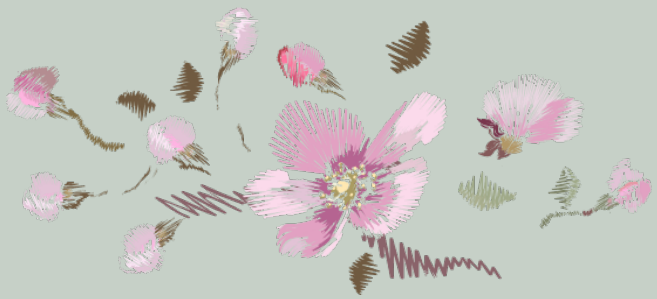
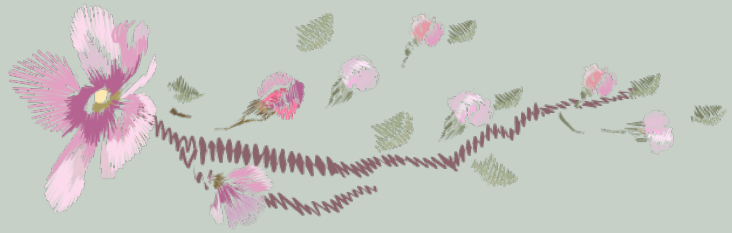
Early articulatory accounts of the syllable claimed that each syllable was the result of a **contraction of the respiratory muscles**.
However, subsequent work showed that there was no direct relationship between syllables and muscle activity: certain bisyllabic (two syllable) words were shown to involve only a single contraction of the expiratory muscles, and, conversely, certain monosyllabic (one-syllable) words were produced with two muscular contractions.. A theory of this kind was put forward by the psychologist R. H. Stetson, who suggested that every syllable is initiated by a chest pulse, a contraction of the muscles of the rib cage that pushes more air out of the lungs. Unfortunately, subsequent investigations have failed to confirm his theory.

Auditory definitions of the syllable make reference to the notion of **sonority**, the relative loudness of sounds. Even if all speech sounds were produced with the same length, stress, and pitch, some are louder than others. For example, vowels are louder than consonants, and within the class of vowels, [ɑ] is louder than [i]. When speech sounds are ranked in terms of their loudness, a ranking known as **sonority hierarchy** is created



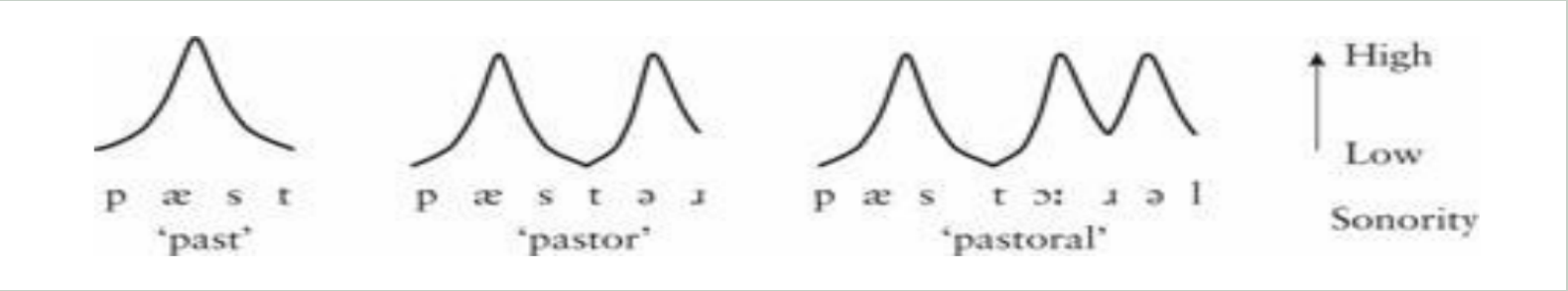
The two primary contributors to the perception of sonority are degree of **opening of the vocal tract** (the more open the tract, the louder the sound is) and **voicing**. Going from most to least sonorous, the order in this hierarchy is: **vowels → approximants → nasals → fricatives → plosives**.

Additional fine-grained distinctions can be made within each **manner class**, such as low vowels and voiced plosives being more sonorous than high vowels and voiceless plosives, respectively


Phonologists have proposed that each syllable is associated with a **peak in sonority**.

The sonority principle stipulates that sonority increases from syllable onset, reaching a maximum at syllable nucleus, and then decreases toward syllable coda. Words such as past, pastor, and pastoral all fit the pattern predicted by the sonority principle, with one, two, and three sonority peaks.



Unfortunately, some English words violate the principle, most notably those beginning with [s]+plosive clusters such as *speak* [spi:k]. Since [s] is judged as more sonorous than all other voiceless and voiced plosives, a word such as *speak* would have two sonority peaks, yet it consists of only a single syllable.

It obviously fails in a word such as *spa*. This word is one syllable, but it must be said to contain two peaks of sonority. It consists of three segments, the first and last of which have greater sonority than the second.

A sonority theory also **fails** to account for the difference in the number of syllables in the phrases *hidden aims* and *hid names*. For speakers who do not have a second vowel in *hidden*, each of these phrases may contain the same sequence of segments, namely, [hɪdneɪmz]; therefore, there are the same number of peaks of sonority. But the first phrase has three syllables, and the second has two.

In all these cases, a sonority theory of the syllable is inadequate: *paddling, frightening, reddening*. Each of these words can be said as two syllables, or they can be said as three syllables, with a syllabic nasal or lateral in the middle. The variations in the number of syllables cannot be said to be due to variations in the number of peaks of sonority.

Another theory is that syllables are marked not by peaks in sonority but by **peaks in prominence**. The relative prominence of two sounds depends in part on what their relative sonority would have been if they had had the same length, stress, and pitch, but it also depends in part on their actual stress, length, and pitch. Then we can say that the [n] in *hidden aims* constitutes a peak of prominence because it has more stress or more length (or both) than the [n] in *hid names*.

The problem with this kind of definition is that
1. one cannot state a cross-linguistically valid procedure for combining sonority, length, stress, and pitch so as to form prominence.
2. part of the problem is that the perceived prominence of sounds relies on language-specific weighting of phonetic factors such as length and sonority, and
3. another part of the problem is that what makes a sound prominent is its position in a word. There is, thus, no way in which one can measure the prominence of a sound.

As a result, the notion of a peak of prominence becomes a completely subjective affair—it does not really shed any light on how one defines a syllable.

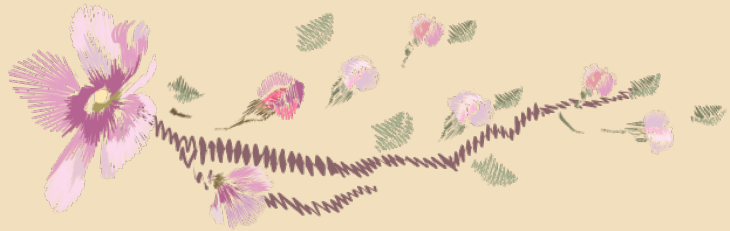
Cross-linguistic variation in suprasegmental features

Languages differ considerably in the syllable structures that they permit. As we have noted;

English has complex onsets and codas. However, consonant clusters within the onset and coda are not randomly arranged, there are rules that govern these sequences.

Hawaiian allows no more than one consonant in an onset, and none in the coda, so that every word (e.g., Honolulu and Waikiki) ends in a vowel.

Standard Chinese allows only nasal consonants in the coda, producing words such as Beijing and Shanghai.


Stress

Stress is a suprasegmental feature of utterances. It applies not to individual vowels and consonants, but to whole syllables. A stressed syllable is pronounced with a **greater amount of energy** than an unstressed syllable and is more **prominent** in the flow of speech.

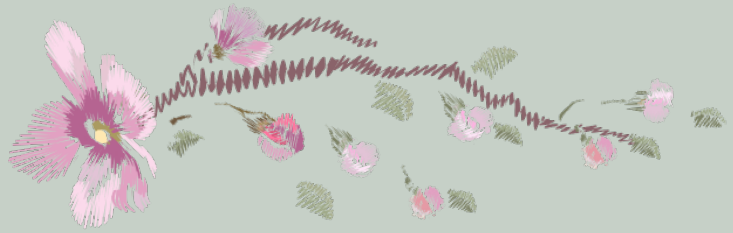
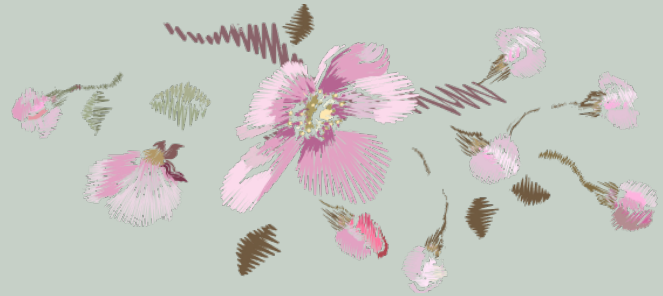
There are three levels of stress: **primary stress, secondary stress,**& **unstressed**.

Articulatorily, a stressed syllable is usually produced with greater physical effort, or greater respiratory activity.

Acoustically, the four main acoustic correlates of stress are **fundamental frequency**, **duration**, **intensity**, and **formant frequency pattern**. Perceptually, the former characteristics correspond to: pitch, length, loudness, and quality. Compared to an unstressed syllable, a stressed syllable in English typically has a higher fundamental frequency, greater duration, greater intensity, and a less reduced (“schwa-like”) quality..

It is sometimes observed that rather than an increase per se in fundamental frequency, a **change** in any direction is a correlate of stress. That is, a stressed syllable could have either a higher or lower fundamental frequency than its unstressed counterpart.

Fundamental frequency and **duration** may be stronger cues to stress than **intensity**. Fundamental frequency signals tone and intonation, duration is important in languages with phonemic distinctions in length, and **formant frequency pattern** is the primary correlate of vowel quality. The relative importance of a particular acoustic correlate of stress may in part be determined by its role in signaling other distinctions.

Stress placement in individual words is governed by specific rules based on the grammatical category, the morphological structure, the number of syllables, and the phonological make up of these syllables.
Function words (articles, conjunctions, prepositions, and pronouns) are usually used in their **weak forms**.(unless we want to stress them for a reason).
However, **content words** (such as nouns and verbs) are usually stressed.

At the sentence level, the stress of a word can change as a function of the stress of neighboring words. This phenomenon is known as **stress shift** or **rhythm reversal**. Depending on the structure of the sentence, a given word in a sentence can be pronounced with or without stress..

When a word occurs as part of a sentence, a **pitch accent** or **nuclear accent** can be used to make certain stressed syllables more prominent than others. A pitch accent is any pitch configuration that serves to make a specific syllable in an utterance relatively more prominent.

In **English** and other stress-timed languages, the accented syllable is always a stressed syllable. The last accented syllable of an intonational phrase is also more prominent than any other accented syllable and is therefore said to have a special type of pitch accent known as nuclear accent.
The location of the nuclear accent is determined by **discourse properties** that the speaker wants to highlight for the hearer (e.g., whether new information is brought into a conversation, or whether two items are contrasted).

Foot and Rhythm

In addition to the syllable, the **foot** is also an important constituent in any discussion of stress. A foot can be defined as consisting of the stressed syllable and all unstressed syllables that follow it. It is claimed that stressed syllables alternate with unstressed syllables to create a **rhythm** of prominence peaks. In a word such as information [ˌɪn.fə.ˈmeɪ.ʃn̩] the stressed syllables [ɪn] and [meɪ] are each followed by an unstressed syllable. Thus, the word information consists of two feet, [ɪn.fə] and [meɪ.ʃn̩].

Stress tends to occur at regular intervals. As a general rule, English does not have stresses too close together. This tendency may cause the stress on a polysyllabic word to be on one syllable in one sentence and on another syllable in another sentence.

Cross-linguistic variation in suprasegmental features

Languages in which more and less prominent syllables alternate in this way are known as **stress-timed** languages. Within the group of stress-timed languages, a distinction is made between **fixed** and **variable** stress.

In a language with fixed stress, primary stress always falls on the same syllable. For example, in **Czech** stress falls on the **first syllable**, in **Polish** and **Swahili** words are almost always stressed on the **penultimate syllable**

In variable stress-timed languages such as **English** or **German**, stress does not consistently fall on a fixed syllable in a word; it is variable across words. It is claimed that in stress-timed languages like English and other Germanic languages, linguistic rhythm is based on the **isochrony** of interstress intervals. That is, in these languages, interstress intervals tend to have the same duration.

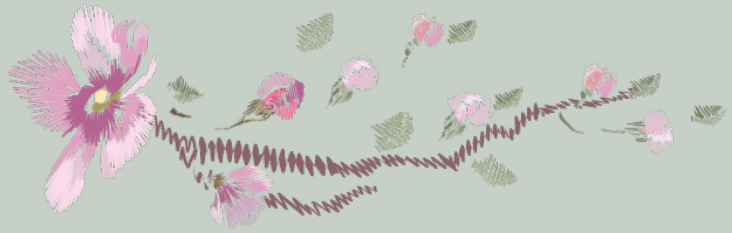
Cross-linguistic variation in suprasegmental features

In **syllable-timed** languages like **French** and **Spanish**, all syllables, stressed and unstressed, are assumed to occur at regular intervals, in contrast to stress-timed languages (in which stressed syllables are claimed to occur at regular intervals). Linguistic rhythm is thought to arise from the isochrony of syllables. In other words, for syllable-timed languages, stressed and unstressed syllables do not differ significantly in duration.

Japanese may provide an example of a third type of timing, **mora timing**. Japanese has a unit called the mora. In Japanese, the difference between a syllable and a mora is that a consonant in the coda of a syllable forms a mora of its own (and so does the second half of a long vowel). According to traditional accounts, each mora takes up approximately the same duration

Length

The acoustic correlate of length is **duration**. Duration is affected by many factors ranging from the segmental to the sentential level. Articulatorily, duration is determined by the timing of articulatory movements. The duration of the segment, syllable, foot, and sentence is affected by a number of factors. One factor that affects duration strongly is the presence of a phonemic length distinction, known as **contrastive length**. Many languages contrast long and short vowels. Duration seems to be the primary cue to vowel length in languages such as **Arabic**, **Danish**, **Estonian**, **Finnish**, and **Japanese**. Languages such as **Dutch**, **English**, and **Swedish**, however, the distinction is primarily one in terms of quality and only secondarily in terms of duration.


Although less common than vowel length contrasts, contrasts between long and short consonants also occur in a variety of languages. **Italian** is an example with minimal pairs such as fato [ˈfato] ‘fate’ and fatto [ˈfatːo] ‘done.’ These long consonants are known as **geminates**. Other well documented languages with a distinction between single and geminate plosives include **Arabic**, **Bengali**, **Estonian**, **Finnish**, **Swiss German**, and **Turkish**, with long consonants generally being two to three times as long as their short counterparts.

There are additional local conditioning factors for segment duration. One such factor was that vowels are longer when followed by a **voiced** as compared to a **voiceless** consonant.

Segment duration is shortened as the **number** of segments and syllables increases. For example, the duration of the vowel [u] will be shortest in the word *fruitier*, somewhat longer in *fruity* and longest in *fruit*.

A syllable is substantially longer at the end of a phrase than at the beginning or middle of a phrase, a phenomenon known as **phrase-final lengthening.**

Tone and Intonation

The **pitch** of a sound is an auditory property that enables a listener to place it on a scale going from low to high. Pitch is the perceptual correlate of **fundamental frequency**, which makes **vocal folds vibration** its articulatory correlate.

The pitch of the voice usually indicates whether the speaker is male or female (**gender**) and what his or her **age** is. In addition, it conveys a great deal of **nonlinguistic information** about **the speaker's emotional state**—whether the person is calm or angry, or happy or sad.

All languages use pitch to mark the boundaries of syntactic units. In nearly all languages, the completion of a grammatical unit such as a normal sentence is signaled by a **falling pitch**. Conversely, incomplete utterances, such as mid-sentence clause breaks where the speaker intends to show that there is something still to come, often have a basically **rising intonation**.

In addition to syntactic and discourse functions, in many other languages, pitch variations can be used to signal **the meaning of a word**. In **Mandarin Chinese**, for example, the consonant–vowel sequence [ma] pronounced with a high and level pitch means ‘mother’, but the same sequence pronounced with a high falling pitch means ‘scold’.

Pitch variations that affect the lexical or grammatical meaning of a word are called **tones**. In the majority of the languages in the world, the meaning of a word depends on its tone.

A language in which the meaning of a word depends on its tone is known as a **tone language**. Tone languages can have either **register tones** or **contour tones**.

In a register tone language, the relative height of each syllable's pitch within the speaker's pitch range (or register) provides a cue to word meaning. These languages typically have up to four distinctive pitch levels. The vast majority of languages spoken in Africa are register tone languages (including **Igbo**, **Shona**, **Yoruba**, and **Zulu**).

In a contour tone language, pitch movement, instead of pitch level, serves to distinguish word meaning. Many of the languages spoken in Southeast Asia (including **Mandarin Chinese**, **Thai**, and **Vietnamese**) are contour tone languages.

The “**tones**” in an English sentence do not affect the meaning of the individual words, although they may affect the meaning of the phrase or sentence.

English has **stress contrasts**, but not tone contrasts.

Japanese is a more striking case of a language that is in some ways between a tone language and a stress language. Words in Japanese have an accent on a particular syllable in much the same way that English words have one or more stresses. In Japanese, the accent is invariably realized as a high pitch, so that Japanese is often called a **pitch-accent** language.

Intonation is the distinctive use of pitch over units larger than a single word. Intonation may convey **linguistic information** by marking the boundaries of syntactic units or by distinguishing different sentence types such as statements, questions, and commands.

Intonation may also convey **nonlinguistic information** such as boredom, impatience, or politeness. Many languages indicate the end of a syntactic unit such as a phrase or sentence by means of a **falling pitch** (in declarative sentences). Many languages mark a yes–no question or other requests for information with a **rising intonation**. Although English declarative sentences and questions usually differ in terms of word order (This is Jane. vs. Is this Jane?), the same declarative sentence could be a statement or question, depending on its intonation.

The declarative sentence is marked by a general lowering of the fundamental frequency throughout the sentence, a phenomenon known as **declination**. In contrast, for the yes–no question a steep rise is evident, particularly at the end of the sentence.

References

Ladefoged, P. and Johnson, K. (2015) (7th edition). *A Course in Phonetics*. Australia: Thompson Wadsworth.

Reetz, H. and Jongman, A. (2009). *Phonetics: Transcription, Production, Acoustics and Perception*. United Kingdom: Wiley-Blackwell.

From Rules to Constraints: Addressing Drawbacks in Generative Phonology

- A Comparative Overview with Examples
- Yasamin Mudhar (2025)

Introduction

- - Generative Phonology (SPE): Introduced by Chomsky & Halle (1968)
- - Key features: Underlying representations, ordered rules, derivations
- - While revolutionary, it faced criticisms leading to new models

Drawback 1 – Overgeneration

- - SPE rules can describe unattested and implausible processes
- - Solution in OT: Universal and violable constraints limit outputs

- English: OT prevents forms like [ngata] for 'data'
- Arabic: Prevents sequences like [ktlb]

Drawback 2 – Opacity & Rule Ordering

- - Complex rule interaction often yields opaque surface forms
- - Natural Phonology: Eliminates rule ordering
- - Harmonic Serialism: Transparent derivation

- English: /divine/ → [dɪˈvaɪn], /divinity/ → [dɪˈvɪnəti]
- Arabic: /yaktubu/ → [yiktib]

Drawback 3 – Unnatural Rules

- - SPE allows implausible derivations
- - Natural Phonology: Only "natural" processes
- - OT: Only outputs satisfying constraints survive

- Arabic: Avoidance of CCC clusters—/ktb/ → [katab]

Drawback 4 – Lack of Universality

- - SPE rules are language-specific
- - OT: Universal constraint set (CON)

- Arabic: Preference for CV syllables
- English: Complex onsets like [str] allowed by different ranking

Drawback 5 – Abstractness

- - URs often lack psycholinguistic evidence
- - Natural Phonology: Focus on perceptible patterns
- - OT: Surface-based evaluation

- English: /bæt/ for 'bat' vs. [bæt]
- Arabic: [ʔamara] instead of abstract /ʔmr/

Drawback 6 – Surface Well-Formedness

- - SPE ignores phonotactic acceptability
- - OT: Direct evaluation of surface forms

- English: [blæk] OK, [bnɪk] *
- Arabic: /ktb/ → [katab], not *[ktb]

Drawback 7 – Prosody & Suprasegmentals

- - SPE poorly models stress, tone, syllable weight
- - Metrical & Autosegmental Phonology improve modeling

- English: 'blackbird vs. black 'bird
- Arabic: /katab/ → ['ka.tab], /maktab/ → ['mak.tab]

Drawback 8 – Weak Phonetic Integration

- - SPE separates phonology from phonetics
- - Natural & OT Phonology ground processes in phonetics

- English: /t/ → [ɾ] in 'butter'
- Arabic: Emphatic spreading from /ṣ/

Summary Table

- Drawbacks & Solutions:
- - Overgeneration → OT constraints
- - Opacity → NP, Harmonic Serialism
- - Unnatural → NP, OT
- - Lack of universality → OT
- - Abstractness → NP, OT
- - Surface well-formedness → OT
- - Prosody → MP, AP
- - Phonetic integration → NP, OT

Thank You

- Questions?

- Further reading:
- - Chomsky & Halle (1968)
- - Prince & Smolensky (1993)
- - Goldsmith (1976)
- - Kiparsky (1982)