

PROSODY

Definition:

In linguistics, **prosody** (from Ancient Greek προσωδία *prosōidía* [proso:(i)día:], "song sung to music; tone or accent of a syllable") is the **rhythm**, **stress**, and **intonation** of **speech**. Prosody may reflect various features of the speaker or the utterance: the emotional state of the speaker; the form of the utterance (statement, question, or command); the presence of **irony** or **sarcasm**; emphasis, **contrast**, and **focus**; or other elements of language that may not be encoded by **grammar** or by choice of **vocabulary**.

Classification:

Languages can be classified according to the distinctive prosodic unit that gives a language its rhythm. Languages can be stress-timed, **syllable**-timed, or **mora**-timed. Stress-timed languages include **English** and **Dutch**, syllable-timed languages include **Spanish** and **Italian**, and mora-timed languages include **Japanese**. The classification of languages is done under the assumption that a language has "isochronous rhythm",^[1] meaning that there is an equal amount of time between stressed syllables, symbols, or moras, depending on the category of language.

Acoustic Attributes:

In terms of **acoustics**, the prosodics of **oral languages** involve variation in **syllable weight**, **loudness**, and **pitch**. In **sign languages**, prosody involves the rhythm, length, and tension of gestures, along with mouthing and facial expressions. Prosody is typically absent in writing, which can occasionally result in reader misunderstanding. **Orthographic** conventions to mark or substitute for prosody include **punctuation** (commas, exclamation marks, question marks, **scare quotes**, and **ellipses**), and typographic styling for **emphasis** (italic, bold, and underlined text).

The details of a language's prosody depend upon its **phonology**. For instance, in a language with **phonemic vowel length**, this must be marked separately from prosodic syllable weight. In similar manner, prosodic pitch must not obscure **tone** in a **tonal language** if the result is to be intelligible. Although tonal languages such as **Mandarin** have prosodic pitch variations in the course of a sentence, such variations are long and smooth contours, on which the short and sharp lexical tones are superimposed. If pitch can be compared to ocean waves, the swells are the prosody, and the wind-blown ripples in their surface are the lexical tones, as with stress in English. The word *dessert* has greater stress on the second syllable, compared to the noun *desert*, which has greater stress on the first (in its "arid land" meaning, but not in its "thing which is deserved" meaning); but this distinction is not obscured when the entire word is stressed by a child demanding "Give me dessert!" Vowels in many languages are likewise

pronounced differently (typically less **centrally**) in a careful rhythm or when a word is emphasized, but not so much as to overlap with the **formant** structure of a different vowel. Both **lexical** and prosodic information are encoded in rhythm, loudness, pitch, and vowel formants.

Domain:

Prosodic features are **suprasegmental**. They are not confined to any one **segment**, but occur in some higher level of an utterance. These **prosodic units** are the actual phonetic "spurts", or chunks of speech. They need not correspond to grammatical units such as **phrases** and **clauses**, though they may; and these facts suggest insights into how the brain processes speech.

Prosodic units are marked by phonetic cues. Phonetic cues can include aspects of prosody such as **pitch**, **pauses**, and **accents**, all of which are cues that must be analyzed in context, or in comparison to other aspects of a sentence. Pitch, for example, can change over the course of a sentence. In English, falling intonation indicates a declarative statement while rising intonation indicates an interrogative statement. Pauses are important prosodic units because they can often indicate breaks in a thought and can also sometimes indicate the intended grouping of nouns in a list. Breathing, both inhalation and exhalation, seems to occur only at these pauses where the prosody resets. Prosodic units, along with function words and punctuation, help to mark clause boundaries in speech. Accents, meanwhile, help to distinguish certain aspects of a sentence that may require more attention. English often utilizes a **pitch accent**, or an emphasis on the final word of a sentence. **Focus accents** serve to emphasize a word in a sentence that requires more attention, such as if that word specifically is intended to be a response to a question.^[1]

"Prosodic structure" is important in language contact and **lexical borrowing**. For example, in **Modern Hebrew**, the XiXéX verb-template is much more productive than the XaXáX verb-template because in morphemic adaptations of non-Hebrew stems, the XiXéX verb-template is more likely to retain – in all conjugations throughout the tenses – the prosodic structure (e.g., the consonant clusters and the location of the vowels) of the stem.^[2]

Unique prosodic features have been noted in **infant-directed speech (IDS)** - also known as **baby talk**, **child-directed speech (CDS)**, or **motherese**. Adults, especially caregivers, speaking to young children tend to imitate childlike speech by using higher and more variable pitch, as well as an exaggerated stress. These prosodic characteristics are thought to assist children in acquiring phonemes, segmenting words, and recognizing phrasal boundaries. And though there is no evidence to indicate that infant-directed speech is necessary for language acquisition, these specific prosodic features have been observed in many different languages.^[3]

Prosody is useful for listeners as they perform sentence parsing. Prosody helps resolve sentence ambiguity. For example, the sentence "They invited Bob and Bill and Al got rejected" is ambiguous when written, although addition of a written

comma after either "Bob" or "Bill" will remove the sentence's ambiguity.^[1] But when the sentence is read aloud, prosodic cues like pauses and changes in intonation will make the meaning clear. The prosody of an ambiguous sentence biases a listener's interpretation of that sentence.^[1] Moving the intonational boundary in the above example will change the interpretation of the sentence. This result has been found in studies performed in both English and Bulgarian.^[4] Prosody is also useful in expressing (for speakers) and detecting (for listeners) sarcasm. The most useful prosodic feature in detecting sarcasm is a reduction in the mean fundamental frequency relative to other speech for humor, neutrality, or sincerity. While prosodic cues are important in indicating sarcasm, context clues and shared knowledge are also important.^[5]

Emotion:

Emotional prosody is the expression of feelings using prosodic elements of speech. It was considered by [Charles Darwin](#) in *The Descent of Man* to predate [the evolution of human language](#): "Even monkeys express strong feelings in different tones – anger and impatience by low, – fear and pain by high notes."^[6] [Native speakers](#) listening to actors reading emotionally neutral text while projecting emotions correctly recognized happiness 62% of the time, anger 95%, surprise 91%, sadness 81%, and neutral tone 76%. When a database of this speech was processed by computer, segmental features allowed better than 90% recognition of happiness and anger, while suprasegmental prosodic features allowed only 44%–49% recognition. The reverse was true for surprise, which was recognized only 69% of the time by segmental features and 96% of the time by suprasegmental prosody.^[7] In typical conversation (no actor voice involved), the recognition of emotion may be quite low, of the order of 50%, hampering the complex interrelationship function of speech advocated by some authors.^[8] However, even if emotional expression through prosody cannot always be consciously recognized, tone of voice may continue to have subconscious effects in conversation. This sort of expression stems not from linguistic or [semantic](#) effects, and can thus be isolated from traditional linguistic content. Aptitude of the average person to decode conversational implicature of emotional prosody has been found to be slightly less accurate than traditional facial expression discrimination ability; however, specific ability to decode varies by emotion. These emotional^[clarification needed] have been determined to be ubiquitous across cultures, as they are utilized and understood across cultures. Various emotions, and their general experimental identification rates, are as follows:^[9]

- Anger and sadness: High rate of accurate identification
- Fear and happiness: Medium rate of accurate identification
- Disgust: Poor rate of accurate identification

The prosody of an utterance is used by listeners to guide decisions about the emotional affect of the situation. Whether a person decodes the prosody as positive, negative, or neutral plays a factor in the way a person decodes a facial expression accompanying an utterance. As the facial expression becomes closer

to neutral, the prosodic interpretation influences the interpretation of the facial expression. A study by Marc D. Pell revealed that 600 ms of prosodic information is necessary for listeners to be able to identify the affective tone of the utterance. At lengths below this, there was not enough information for listeners to process the emotional context of the utterance.^[10]

An [aprosodia](#) is an acquired or developmental impairment in comprehending or generating the emotion conveyed in spoken language. Aprosody is often accompanied by the inability to properly utilize variations in speech, particularly with deficits in ability to accurately modulate pitch, loudness, intonation, and rhythm of word formation.^[11] This is seen sometimes in persons with [Asperger syndrome](#).^[12]

Brain Regions Involved:

Producing these nonverbal elements requires intact motor areas of the face, mouth, tongue, and throat. This area is associated with [Brodmann areas 44 and 45 \(Broca's area\)](#) of the left [frontal lobe](#). Damage to areas 44/45 produces motor aprosodia, with the nonverbal elements of speech being disturbed (facial expression, tone, rhythm of voice).

Understanding these nonverbal elements requires an intact and properly functioning right-hemisphere [perisylvian area](#), particularly [Brodmann area 22](#) (not to be confused with the corresponding area in the *left* hemisphere, which contains [Wernicke's area](#)).^[13] Damage to the right inferior frontal gyrus causes a diminished ability to convey emotion or emphasis by voice or gesture, and damage to right superior temporal gyrus causes problems comprehending emotion or emphasis in the voice or gestures of others. The right Brodmann area 22 aids in the interpretation of prosody, and damage causes sensory aprosodia, with the patient unable to comprehend changes in voice and [body language](#).

