In Gurevich, N. & Grindrod, C. (Eds.) 2022. *Clinical Applications of Linguistics to Speech-Language Pathology.* Taylor & Francis.

Prosody: Linguistic and clinical perspectives

Jennifer Cole[1], Allison Hilger[2], Shivani Patel[3]
[1]*Northwestern University,* [2]*University of Colorado Boulder,* [3]*Lucile Packard Children's Hospital Stanford*

Abstract. Prosody is sometimes described as the musical quality of speech, and though it involves melodic and rhythmic dimensions of sound that are also primary characteristics of music, it is far from ornamental and has critical roles signaling linguistic structure in continuous speech, and conveying meaning related to a speaker's communicative intentions. This chapter reviews the phonological and phonetic characteristics of prosody in relation to their linguistic function, with reference to English, and the physiological requirements for the expression of prosody in speech, and then reviews some of the more common patterns of atypical prosody associated with structural and/or neurological impairment in adult special populations. The chapter highlights methods for assessing prosody and current approaches to treatment, and discusses some of the consequences of atypical prosody for social interaction.

As a clinician, you may often note disrupted or atypical prosody in the presence of structural or neurological impairment in adults, impacting linguistic communication and social interaction. But *what exactly are the acoustic and articulatory characteristics of prosody, and why is it that prosodic impairments can disrupt communication?* This chapter addresses these questions from the perspective of current linguistic theory and introduces current approaches to the assessment and treatment of deficits in the production of prosody. A discussion of the key role of prosody in successful communication and social interaction highlight the importance of addressing prosody in the clinical domain.

1. Introduction

Prosody refers to the patterning of pitch, timing, loudness, voice quality, and timbre across the syllables of a spoken word, phrase, sentence, or larger discourse unit. Prosodic patterning is rich in its potential to convey *linguistic* meaning related to the interpretation of words, sentences and discourse, and *paralinguistic* meaning related to the speaker's psychological state, social identity, and the communicative setting. Prosody is inherent to speech and a property of every spoken language. At the same time, languages differ from one another in their prosodic systems, with perceptually salient differences in the melodic and rhythmic aspects of speech, and in the mapping from prosodic expressions to linguistic meaning. The distinct prosodic patterns used in a given language, and even in a particular dialect, are part of what a child learns in the normal course of development. Deviations from community norms in the production of prosodic patterns, or in the perception and interpretation of those patterns, can impact communication outcomes.

This chapter presents an overview of prosody as expressed in typical speaker-hearer populations, considering the kinds of linguistic and paralinguistic information that prosody conveys (Section 1), the phonological and phonetic encoding of prosody (Section 2), and the physiological mechanisms involved in the production of prosody (Section 3). These topics provide the foundation for the analysis of atypical prosody and prosodic adaptation in special adult populations (Section 4), and for the development of approaches to prosodic assessment and interventions (Section 5). We conclude with a discussion of prosody as it relates more broadly to interactive social behavior (Section 6). Throughout the chapter, reference is made to the prosodic patterns of American English, some of which generalize to other dialects of English. But it is important to bear in mind that prosodic patterns and their function in conveying linguistic meaning are known to differ across languages, and therefore identifying typical or atypical prosody must be done with reference to the norms of a specific speech community.

2. The functions of prosody in conveying linguistic and non-linguistic information

2.1. Prosodic encoding of structure in word, sentence and discourse level juncture

A core function of prosody is the encoding of linguistic structure, on analogy with the function of punctuation, font emphasis (italic, bold), space between words, and paragraph indentation in written language. In some languages, prosodic patterns operating at the word level function to mark word boundaries, for example, through stress assigned to the syllable that sits at or near the initial or final word boundary. English has an especially complex system of word stress, with stress placed at or near the right edge of a word, but subject to numerous constraints on syllable and morphological structure (Fudge, 2015). It's important to note that not all languages distinguish stressed and unstressed syllables within words, and among those that do not are many of the so-called *tone* or *pitch-accent* languages, in which tone features are part of a phonological make-up of the word (i.e., along with its consonants and vowels, as part of its dictionary specification), e.g., Mandarin, Thai, Vietnamese (Gussenhoven, 2004). Yet other languages appear not to have any system of prosodic marking at the word level, e.g., Indonesian (Gordon, 2014).

While word-level prosody is not universal among languages, to the best of our knowledge, all languages use prosody to mark structure at the phrase level (1), and to identify structural relations among successive phrases in complex sentences (2, from Ladd, 2008).

(1) [*When Lily awakened*] [*the baby was crying*]

(2) i. [[A and B] but C]] *"Warren is a stronger campaigner, and Ryan has more popular policies, but Allen has a lot more money."*

ii. [A [but B and C]] *"Warren is a stronger campaigner, but Ryan has more popular policies, and Allen has a lot more money."*

Prosodic phrases are composed in relation to syntactic structure, though they are not necessarily direct extensions of syntactic phrases (Shattuck-Hufnagel & Turk, 1996). Typically, a prosodic juncture (the boundary between successive prosodic phrases) must be located at a syntactic boundary (e.g., a noun phrase, verb phrase, prepositional phrase, or clause), but a single prosodic phrase may span two or more syntactic phrases, e.g., including the subject noun phrase and the verb phrase in (3). It is also possible for a single syntactic phrase to be split into two or more prosodic phrases, as with the more complex verb phrase in (4). Prosodic phrasing reflects speech production planning (Krivokapić, 2014), such that the location of a prosodic phrase boundary between two words depends on the presence of a syntactic phrase boundary at that location, and on the length and complexity of the syntactic constituents preceding and following the syntactic juncture (Watson & Gibson, 2004). Moreover, although syntax constrains prosodic phrasing, speakers have substantial flexibility in the prosodic phrasing of a sentence, both in the number and extent of prosodic phrases, especially for complex or long sentences; non-grammatical factors such as speech rate and style also play a role (Gee & Grosjean, 1983).

(3) [Sam gave the man a ticket]
(4) [Sam gave a certificate of achievement] [to each child who completed the course]

Generally speaking, the presence of a prosodic phrase boundary is a fairly reliable cue for the presence of a syntactic constituent boundary at the same location. Indeed, a prosodic boundary located in a position that is not also a major syntactic boundary, as in (5ii), disrupts sentence processing, and in some instances, may result in faulty sentence interpretation, e.g., (6ii) (Speer et al., 1996)

(5) i. [George and Mary] [gave blood]
ii. ?? [George] [and Mary gave blood]
(6) i. [Whenever the guard checks] [the door is locked]
ii. ?? [Whenever the guard checks the door] [is locked]

Looking above the level of the sentence, prosody further serves to convey dialogue structure in interactive speech. For instance, prosody functions as a resource for managing turn-taking in a conversation, and there are also distinct prosodic patterns used to signal the initiation, end, or continuation of topics (Ward, 2019). When the prosodic marking of the end of a conversational turn is not present, or if it occurs in other locations, it disrupts conversational flow. Such disruptions often result in increased talker overlap, or awkwardly long pauses before a change of talker turn.

2.2. Prosodic prominence

In addition to their grouping function, prosodic phrases also define a domain for prominence, a feature that identifies one or more elements within a prosodic domain as standing out relative to other elements. What it means to “stand out” depends on the level of prosodic structure under consideration. For languages with word-level prosody, word-level prominence corresponds to primary word-level stress, which is located relative to the left or right edge of the word according to language-specific constraints. In some languages, including English, additional syllables in a longer polysyllabic word may be designated as having a lower degree of prominence, realized as secondary stress. The location of primary word-level stress in English is important for accurate word recognition for native (English-speaking) listeners (Cutler & Clifton, 1984).

Prominence is also a feature of prosodic phrases, wherein one word in the phrase is assigned the primary phrasal stress, also termed *nuclear stress* or *nuclear prominence*. In languages with word-level stress, including English, phrasal prominence must be realized on a word that has word-level stress, which excludes stressless monosyllable ‘function’ words (determiners, pronouns, prepositions, conjunctions). In these languages, a syllable that is marked for both word-level and phrasal stress has greater prominence than a syllable that has only word-level stress. In a parallel manner with word-level stress, phrasal stress is located relative to the edge of the prosodic phrase. In English, the rightmost stressable word is assigned phrasal prominence, as illustrated in (7) (the syllable with phrasal and word-level stress is marked by CAPS). In addition to the primary phrasal stress, English also allows for optional secondary prominences on stressable words in *prenuclear* position (preceding the nuclear stress), which are particularly common in phrase-initial position, as illustrated in (8), with optional prenuclear prominences on *Sam* and *usually*, and nuclear phrasal prominence on *tickets*).

(7) i. [Sam gave the man the TICKet]

ii. [Sam gave the TICKet to him]

(8) [SAM/Sam was USuallly/usually the one to buy the **TICK**ets]

Because every stressable word in a sentence is eligible to realize phrasal prominence, the occurrence of a stressable word with phrasal stress, by itself, does not provide information about sentence structure. All possible assignments of phrasal prominence to stressable words are prosodically well-formed. But as we will see in the next section, in English the location of the primary phrasal stress encodes *information structure* distinctions related to focus and the status of a word as introducing 'new' or 'given' information relative to the discourse context, and a mismatch between phrasal prominence and the information structure context can impair sentence comprehension.

2.3. Pragmatic meaning

Beyond its structure-marking function, in English prosody also conveys pragmatic meaning (i.e., meaning related to the context of an utterance) through intonation, which is the specification of tone features associated with prominent words and the edges of prosodic phrases. English is not alone in conveying pragmatic meaning through intonation, but the richness of the system is a hallmark of English and other West Germanic languages. Two distinct pragmatic functions are encoded through intonation in English: information structure and speech act meaning. Information structure relates to *reference*—the entity (physical or abstract) in the world that a linguistic expression refers to—and covers notions like corrective or contrastive focus (9) and the status of a referent as discourse-new or discourse-given (10) (Brazil, 1980; Hirschberg, 2015; Westera et al., 2020). English marks information structure distinctions through the phonetic implementation of phrasal stress and in the choice of tone features (see Section 2) associated with words that have corrective focus (the bolded word in 9i) or contrastive focus (9ii), words that convey the answer to a question (i.e., narrow focus, 9iii), and for words that introduce new information to the discourse (10i). If a discourse-given word occurs in the default (phrase-final) position for nuclear stress (the underlined *bears* in 10ii*),* it will typically not be assigned nuclear stress, which instead shifts leftward to the nearest stressable word (*see*).

(9) i. Speaker A: I think Sam is going to the meeting tomorrow.
Speaker B: No, **Sue** was asked to go instead of Sam.

ii. Speaker A: Sue hates to travel so I was surprised she volunteered when Sam was unable to go to the meeting.
Speaker B: Well, Sue was **asked** to go in his place.

iii. Speaker A: Sam had to cancel, so who will go to the meeting tomorrow?
Speaker B: **Sue** was asked to go his place.

(10) i. (The guidebook says wildlife is abundant in this park…) I don't see any **bears**.

ii. (The guidebook says bears live around here….) I don't **see** any <u>bears</u>.

As with prosodic boundaries, listeners also pay attention to phrasal stress in sentence comprehension. In particular, the absence of phrasal stress on the rightmost stressable word in the prosodic phrase is a strong cue that that word is discourse-given. Likewise, the location of primary phrasal stress on a word in non-final position, especially when the final word is discourse-new, leads listeners to interpret the non-final word as focused, as in (9ii, iii). As illustrated above, under typical conditions the prominence status of a word is congruent with the prior discourse context that establishes focus and givenness. But in the event of a mismatch, e.g., where phrasal stress is realized on a phrase-final word that is discourse-given (e.g., if phrasal stress is on "bears" in 10ii), listeners can be confused about the intended referent of the expression. Faced with a speaker who is unreliable in producing prominence patterns that are congruent with the discourse context, listeners rapidly adapt by disregarding phrasal stress as an interpretive cue to focus or givenness (Roettger & Franke, 2019).

Speech act meaning (also called *illocutionary force*) concerns the speaker's communicative goal (e.g., to assert, inquire, request, contradict). English conveys speech act meaning through the choice of sentence type (e.g., declarative, wh-question, imperative) paired with *nuclear tune,* which is the pitch melody spanning the nuclear stress to the end of the intonational phrase. For example, a declarative sentence with a falling tune is the typical way to express an assertion, but the same sentence paired with a rising tune can be used to seek confirmation or to inquire. Contemporary linguistic accounts of English intonation derive these tunes from a sequence of tone features marking the nuclear stress and prosodic phrase boundary (see Section 2) (Gussenhoven, 2004; Ladd, 2008).

2.4. Paralinguistic meaning related to social factors and speaker affect

Evaluating prosody for its function in marking linguistic structure and conveying pragmatic meaning is complicated by the fact that the speech channel simultaneously conveys information about the speaker's affect, in terms of emotion, attitude and mood. There is an expansive literature on affective prosody (also known as *emotional* prosody) showing that variation in pitch, tempo, and other acoustic-prosodic parameters correlates with enacted or perceived distinctions in speaker affect along dimensions such as arousal/potency and valence (Bänziger & Scherer, 2005). To date there is little research examining linguistic and emotional prosody together to understand how they contribute independently or jointly in any given utterance. Yet, since similar acoustic patterns are seen in linguistic and emotional prosodic expression, this is a question that is important for anyone evaluating prosody in a research or clinical setting, and an area in need of further research.

The social context in which an utterance occurs also influences a speaker's choice of prosodic expression. In contexts of interactive communication, social factors may play a role in listeners' interpretation of pragmatic meaning conveyed through prosody. Prosody can index a speaker's social affiliation, illustrated for example in the distinct prosodic patterns associated with ethnic or regional dialects, gender, or sexual orientation (Holliday, 2021). Another dimension of social meaning tied to prosody is politeness, where higher pitch, slower speech rate, and increased vowel (or syllable) duration are associated with perceived politeness (Navarro & Nebot, 2014).

3. Prosody in phonological representations and its phonetic expression

Prosodic structure is part of the phonological representation of words and phrases. At the word level, prosodic structure consists of 'metrical feet' that bundle successive syllables together and which determine the placement of stress. In English, the metrical foot may include two syllables, with stress on the leftmost of the pair; primary stress goes on the rightmost foot in the word, and in longer words a secondary stress is possible on other feet, subject to constraints on rhythmic stress alternation. At the phrase level, there is a hierarchy of prosodic phrases, with smaller *intermediate* phrases that are combined into larger *intonational* phrases, as illustrated in Fig. 1.[1] At each level of prosodic structure, one element is designated as prominent. In the intermediate phrase, prominence is on the rightmost

[1] This model is based on the work of Pierrehumbert (1980), and is couched within Autosegmental-Metrical theory, a phonological theory of prosody, tone and intonation. See Ladd (2008).

stressable word, which marks the location of the nuclear (phrasal) stress (*student* and *daily* in Fig. 1). The perceptually strongest prominence in the intonational phrase is the nuclear stressed word of the rightmost intermediate phrase, *daily* in Fig. 1 (Cole et al., 2010, 2019). The nuclear prominence in the intermediate phrase is marked with an obligatory pitch accent—a low or high tone feature associated with the stressed syllable (annotated L* or H*).[2] There are also optional prenuclear pitch accents on earlier stressed words. The right edge of the intermediate and intonational phrases are marked with tone features (annotated H- or L-, for intermediate phrases, and H% or L% for intonational phrases). These edge-marking tones combined with the preceding nuclear pitch accent together derive the typical intonation pattern of English: a perceptually salient, dynamic pitch movement at the end of a phrase. Since the end of a sentence tends also to be the end of an intermediate phrase and intonational phrase, the nuclear 'tune' in English tends to occur at the end of a sentence. This pattern is overridden in sentences that mark narrow, corrective or contrastive focus on a word located earlier in the sentence (as in 9, above), in which case the nuclear pitch accent is located on the focused word, with loss of prominence ('deaccenting') on any following word.

H* !H* H* Pitch accent
L- L- H% Boundary tone

$_{IP}$[$_{ip}$[The **mu**sic **stu**dents]$_{ip}$ $_{ip}$[practiced **dai**ly]$_{ip}$]$_{IP}$

Figure 1. Prosodic structure and intonational features for a spoken production of the sentence *The music students practiced daily.*

Prosodic structure and associated intonational features (pitch accents, boundary tones) are phonetically expressed through variation in articulation and corresponding variation in acoustic parameters. The details of this phonetic spell-out vary across languages, but generally speaking, prominence is associated with phonetic enhancement, and boundaries are associated with lengthening and pause. These phonetic effects of prosody come about due to changes in laryngeal settings, oral aperture, the volume of airflow across the glottis, and in the timing of speech gestures (e.g., tongue body raising, lip closure, etc.). Articulatory studies of English show that phonetic enhancement associated with prominence is implemented with hyper-articulation (de Jong, 1995; Byrd & Krivokapić, 2021), increased airflow and regular (modal) as opposed to irregular (non-modal) phonation, and laryngeal adjustments involved in the realization of high or low pitch targets in the spell-out of the pitch accents. Different articulatory adjustments are involved in the production of prosodic boundaries. A gradual slowing down of articulatory gestures occurs immediately preceding a prosodic phrase boundary, with maximal effects of lengthening on the final syllable. This is accompanied by hypoarticulation—a reduction in the magnitude of articulatory gestures, especially those in non-prominent syllables. Younger American English speakers increasingly also deploy irregular (creaky) phonation at the end of an intonational phrase, in the presence of low boundary tones.

[2] The inventory of pitch accents in English also includes a downstepped high (!H*), and several accents composed of a two-tone sequence, such as L+H* (Ladd, 2008).

The articulatory effects just described give rise to corresponding acoustic effects of prosody, with prominence manifest in increased intensity and duration, spectral measures (e.g., formants) reflecting gestures at more peripheral locations in the oral cavity (Cho, 2005), and spectral envelope measures of increased vocal effort (Sluijter & van Heuven, 1996). Accompanying these are dynamic patterns of fundamental frequency (f_0) change, producing pitch contours that implement the high and low tones of pitch accents when present. Each of these acoustic effects contribute to making the prominent element stand out perceptually (Cole et al., 2010, 2019). Acoustic correlates of prosodic phrase boundaries are lengthened duration and reduced intensity of the phrase-final syllable (or longer pre-boundary interval), and irregular pitch periods perceived as creaky voice. F_0 correlates of boundaries depend on the tone features associated with the boundary, resulting in a pitch rise, fall, or sustained pitch level from the preceding (nuclear) pitch accent.

4. Physiological requirements for prosody

The control of pitch, loudness, and timing for prosody in speech requires highly precise coordination among the speech subsystems: respiration, phonation, resonance, and articulation (Kent et al., 1989). While accurate prosodic production mainly relies on the subsystems of respiration and phonation, resonance and articulation also play important roles. In this section, we briefly summarize the anatomy and physiology of each subsystem and their roles in the production of prosody.

Respiration is the driving force of speech production (Bunn & Mead, 1971). Intensity and fundamental frequency (f_0) are directly influenced by air pressure from exhalation (Huber, 2008; Titze, 1989). Speech timing is influenced by the location of pauses and breath during speech (Wang et al., 2010). In resting breathing, there is a relatively regular pattern of inspiration and expiration; the diaphragm is the primary muscle involved with little need for accessory muscle involvement. This breathing pattern is vastly modified when speaking (Hixon et al., 1973; Hoit et al., 1989). Speech production requires a long exhale and a short inhale. The exhale must be controlled so that alveolar air pressure can be modified for phonation.

For prosody, speakers make subtle changes in alveolar air pressure to alter vocal intensity (Huber, 2008), for example in a loud restaurant, or to increase intensity at a more local scale for phrasal stress and boundary. Neurological impairment can disrupt respiratory coordination and result in poor loudness control. For example, ataxic dysarthria due to cerebellar damage is characterized by explosive loudness bursts which result from the lack of control over alveolar air pressure during exhalation (Kent et al., 2000). Hypokinetic dysarthria due to Parkinson's disease is also associated with poor loudness control, characterized by reduced loudness with little variation in intensity during speech due to multiple factors including air leakage through bowed vocal folds, lack of control over exhalation, speaking into the expiratory reserve volume, and inhaling to an inadequate inspiratory volume (Huber & Darling-White, 2017).

Respiratory control is required for appropriate phrasing and timing in speech. The length of a prosodic phrase is dependent on the degree of respiratory support (Russell & Stathopoulos, 1988). Being able to produce a longer and slower exhale, for example, allows a speaker to produce more words per phrase before needing to pause to take a breath. When respiratory support is reduced due to physiological or neurological injury, speakers often need to take a breath at every pause and to pause more frequently (Huber, 2008). Accordingly, pauses frequently occur at unusual locations in speech. These atypical pauses can be perceived as unintended phrasal boundaries, leading to shorter prosodic phrases and misinterpretations of prosodic meaning.

Although f_0, the acoustic correlate of pitch, is mainly controlled at the laryngeal level, respiration also plays an important role. Indeed, control of f_0 exemplifies the precise coordination between the phonatory and respiratory systems (Titze, 1989). F_0 is influenced by the degree of subglottal pressure. When respiratory-phonatory coordination is disrupted, there may be inadequate subglottal pressure to effectively set the vocal folds into vibration, leading to reduced vocal quality and poor control of f_0. Greater detail of pitch control will be provided in the next subsection on phonation.

Phonatory control is needed for prosodic aspects of f_0, intensity, and vocal quality. Disruption to laryngeal coordination and respiratory-phonatory coordination is common in neurological injury, resulting in variable pitch and loudness production and poor vocal quality. Changes in vocal fold movement alter the degree of glottal closure and frequency of vibration, leading to perceptual changes in vocal quality, pitch, and, to a lesser extent, intensity.

Although the respiratory and phonatory subsystems are the primary systems of prosodic control, resonance and articulation also play notable roles. Physiological and neurological impairments can impact the control and coordination of the velum, resulting in hypo- or hyper-nasal sounding speech. While nasality in speech does not directly influence the production of vocal intensity and fundamental frequency, it can influence the perception of loudness and pitch. For example, hyper-nasal speech is often perceived as monotone and flat intonation (Tardif et al., 2018).

Similar to resonance, articulation has little direct impact on the control of intensity and fundamental frequency when compared with respiration and phonation, but it does play an important role in timing in speech. The timing of articulatory movements is modified to lengthen or shorten syllable and word duration in speech, which influences the perception of lexical and phrasal stress (de Jong et al., 1993). Stressed syllables are produced with lengthened vowel duration and hyperarticulation. Imprecise articulation, one of the most common characteristics in motor speech disorders, impacts the perception of timing, lexical stress, and phrasal stress.

5. Characteristics of atypical prosody in adult special populations

The production of prosody in speech is accomplished through precise coordination both within and across speech subsystems, which can be disrupted when there is structural and/or neurological impairment. In this section, we describe the characteristics of atypical prosody in dysarthria, apraxia of speech, right hemisphere disorder, and autism spectrum disorder in adults.

5.1. Dysarthria

Dysarthria is defined as an impairment in the execution of speech (Yorkston et al., 1999). It is differentiated from other motor speech impairments involved in the planning and programming of speech (i.e., apraxia of speech). Six subtypes of dysarthria were delineated in the classic 1969 study by Darley, Aronson, and Brown based on collections of perceived articulatory, phonatory, and prosodic errors (Darley et al., 1969). While there is variability within each dysarthria subtype, all subtypes have potential for prosodic impairment, some more directly, such as ataxic dysarthria, and some more indirectly, such as spastic dysarthria.

Ataxic dysarthria is associated with ataxia, or damage to the cerebellum or cerebellar pathways (Kent et al., 2000) and is characterized by both articulatory and prosodic errors. Articulatory errors include irregular consonant production and vowel distortions, and prosodic errors include explosive loudness, variable pitch, a word-by-word cadence, and equal and excess stress patterns. These speech errors result from disruption to the timing, scaling, and coordination of speech movements in the cerebellum (Ackermann, 2008). For example, explosive loudness results from poor control over alveolar air pressure in conjunction with the onset of phonation. The resulting lack of loudness and pitch control create perceptual confusion in identification of pitch accents and phrasal boundaries.

Both hyperkinetic and hypokinetic dysarthrias are associated with basal ganglia impairment (Spencer & Rogers, 2005; Zyski & Weisiger, 1987). The basal ganglia are important neural structures for the control and inhibition of movement, reflected in the differential impairment in hyperkinetic and hypokinetic dysarthria. Hyperkinetic dysarthria is characterized by a disinhibition of movement, resulting in involuntary movement in the body. In speech production, both articulation and prosody are interrupted by abnormal, involuntary movement. Speakers with hyperkinetic dysarthria often speak in short phrases as a compensatory technique to attempt produce a complete message prior to an involuntary movement. The shorter phrasing results in variable rate of speech, strained voice, and reduced phrasal stress. When an involuntary movement occurs during speech, it can cause articulatory breakdown, variations in pitch and loudness, and hypernasality.

Hypokinetic dysarthria, associated with Parkinson’s Disease, is characterized by inhibition of movement, resulting in difficulty in preparing, maintaining, and switching motor programs. In speech, inhibition of movement is reflected in a flat intonation, reduced emotional prosody, and lack of range of motion for articulation. Counterintuitively, the reduction in range of motion results in a rapidly increased rate of speech. Therefore, prosody in hypokinetic dysarthria is exhibited by rapid speech and flat intonation. Because of the lack of variation in pitch, loudness, and syllable duration, perceptual identification of stress and boundary are difficult. Speakers with hypokinetic dysarthria are often misperceived as being bored, angry, or disinterested when the speaker may actually feel much differently (Pell et al., 2006).

Both spastic and unilateral upper motor neuron (UUMN) dysarthrias are associated with damage to the upper motor neuron pathways in the cortex usually due to stroke, tumor, or degenerative disease such as primary lateral sclerosis (Duffy, 2013). Spastic dysarthria is associated with bilateral upper motor neuron damage, whereas UUMN dysarthria is associated with unilateral damage. The bilateral damage in spastic dysarthria typically results in a more severe speech impairment, characterized by slow rate of speech, imprecise consonants, strained vocal quality, spasticity, monopitch, and monoloudness. Prosody in spastic dysarthria will be influenced mostly by the slow rate of speech and lack of control over pitch and loudness. The impairments in pitch, loudness, and vocal quality in spastic dysarthria arise from spasticity in the vocal folds. During phonation, the vocal folds are hyper-adducted, leading to the strained vocal quality and difficulty in modulating pitch and loudness. Prosody is sometimes affected in UUMN dysarthria but usually at a mild severity level, with articulatory breakdown, slow rate of speech, mild hypernasality, and mildly strained phonation. Each of these characteristics can affect the perception of phrasal stress and boundaries in speech. UUMN dysarthria is commonly co-occurring with aphasia and apraxia of speech due to the locus of impairment.

Flaccid dysarthria is the final dysarthria subtype, associated with damage to the lower motor neuron pathways (Duffy, 2013). The site of lesion can occur at four levels within these pathways: the brainstem, cranial or spinal nerves, neuromuscular junction, or muscle. The presentation of flaccid dysarthria has considerable variability depending on the site of lesion. For example, if cranial nerve seven (CN VII) is damaged, the resulting facial paralysis will cause considerable articulatory error. However, if cranial nerve ten (CN X) is damaged, the resulting vocal fold weakness or paralysis will affect vocal quality. Prosody may be impaired in flaccid dysarthria depending on the site of lesion and the severity of damage. If articulation is considerably disrupted, then phrasal stress will be impacted by variable syllable lengthening. If phonation is impaired, then voice quality, pitch, and loudness may become more variable in speech.

5.2. Apraxia of Speech (AOS)

Apraxia of Speech (AOS) is a disorder affecting the ability to plan and program motor plans for speech (Yorkston et al., 1999). Unlike dysarthria, which can occur from damage to multiple parts of the brain, AOS always results from pathology to the left cerebral hemisphere typically from stroke or tumor. The specific neural areas usually affected in AOS are Broca’s Area (i.e., the inferior frontal gyrus) and the Supplemental Motor Area. AOS commonly co-occurs with aphasia (usually Broca’s Aphasia) and UUMN dysarthria.

The planning and programming impairments in AOS cause both articulatory and prosodic disruption. For articulation, consonant and vowel distortions and substitutions are common. These articulatory errors can impact the perception of lexical and phrasal stress because of disruption in syllable lengthening for stress. Prosodic disruption in AOS includes slow rate, syllable segregation, and error in stress assignment. Pitch and loudness control are only affected in AOS in the context of speech production. Individuals with AOS are able to sing with appropriate pitch and loudness control. Because AOS is a disorder of planning and programming speech, and not a disorder of the execution of speech, the impairments are isolated to the context of speech. Therefore, pitch, loudness, and timing are only affected when speaking and not during other tasks involving vocalization.

5.3. Right Hemisphere Disorder (Aprosodia)

Aprosodia, the lack of prosody (rather than disrupted prosody described in dysarthria and AOS), can result from damage to the right hemisphere (Duffy, 2013). In aprosodia, prosody is flat and robotic-like, with significant monotony in pitch, loudness, and duration. Damage to the right hemisphere can cause aprosodia because of the role of the right hemisphere in emotional and affective components of behavior. Therefore, right hemisphere disorder results in a lack of expressive or affective prosody in speech. Their lack of affective prosody is generally not reflective of their true emotional state.

5.4. Autism Spectrum Disorder (ASD) in Adults

While autism spectrum disorder (ASD) is primarily characterized by deficits in social communication and restricted/repetitive behaviors (American Psychiatric Association, 2013), deficits in prosody are among the first detectable characteristics to create an impression of “oddness” among typically developing individuals (Mesibov, 1992; Van Bourgondien & Woods, 1992). Prosodic differences observed in adults with ASD generally impact a range of characteristics, including intonation, stress patterns, speech rate, affective quality, and loudness (Baltaxe et al., 1984; Baltaxe & Simmons, 1985; Baron-Cohen & Staunton, 1994; Fay, 1969; Patel et al., 2020; Pronovost et al., 1966; Shriberg et al., 2001). Notably, there is substantial variation in the prosodic patterns of individuals with ASD, such that differences in any of the observed characteristics (e.g., intonation) may be described as being too flat/monotone for a subset of individuals yet too variable/sing-songy for yet another subset. Adults with ASD show reduced affective prosody recognition (Globerson et al., 2015). An important point is that while commonly noted among individuals with ASD, prosodic differences are not a defining characteristic of the disorder, meaning that a subset of individuals with ASD may present with fairly typical receptive and expressive prosody. Preliminary research suggests that atypical sensory-motor integration, particularly between the auditory (sensory) and vocal (motor) systems contributes to atypical expressive prosody among individuals with ASD (Patel et al., 2019).

6. Linguistically Informed Approaches to Assessment and Treatment

Assessment and treatment of prosody in adults is a difficult task (Peppé, 2009). Prosody is highly variable among individual speakers and norms for typical prosody have not yet been objectively delineated. In this section, we propose strategies for assessment and treatment of prosody based on linguistic theory and physiological knowledge of the speech impairments.

6.1. Assessment

Assessment of prosodic impairment in adults should include estimates of the severity of the impairment, the subsystems that could be contributing to the impairment, and the accuracy and appropriateness of various aspects of prosody. While there are some formal assessments of prosody, issues with validity and efficiency limit the use of these assessments (Peppé, 2009). Current standardized assessments which include prosody are often used mainly with Right Hemisphere Disorder. Some of these assessments include the Burns Brief Inventory of Communication and Cognition: Right Hemisphere Inventory (BBI-RHI; Burns 1997), Mini Inventory of Right Brain Injury (MIRBI-2, Pimental & Knight, 2000), Montreal Protocol for the Evaluation of Communication (Protocol MEC; Joanette et al., 2015), and the RIC Evaluation of Communication Problems in Right Hemisphere Dysfunction (RICE-2; Halper et al., 1991). However, there are norms for typical prosody have not yet been objectively established. Therefore, we propose the use of a combination of perceptual and acoustic measures to determine the severity and features of the prosodic impairment. Detailing these characteristics will allow clinicians to set up goals for treatment.

The first goal in assessment will be to collect observations of which prosodic features are disrupted. Prosodic features to observe include the following:

- Rate of speech
- Length of phrasing
- Location of pauses
- Breathing patterns during speech
- Overall loudness
- **Perceptual salience of word-level and phrasal stress**

- **Perceptual salience of prosodic phrase boundary**
- **Expression of emotion during speech**

We focus on the bolded items in the above list features and how to use linguistically informed assessment to examine prosodic impairment.

Phrasal stress is perceptually salient when there is a change in pitch, increased loudness, lengthened word duration, and hyperarticulation. The clinician should listen to the client's speech in sentence production, passage reading, and spontaneous speech to identify stressed words. When prosody is highly variable, such as in ataxic or hyperkinetic dysarthria, it can be difficult to identify the stressed word when pitch, loudness, and timing vary considerable during the production of each word. Likewise, when prosody has too little variation, such as in hypokinetic dysarthria or right hemisphere disorder, there is little information in pitch, loudness, and timing to identify which word is stressed. In addition to perceptual evaluation, visual inspection of spectrograms and pitch contours in the acoustic speech signals are used to examine patterns of pitch and loudness variation. Visual inspection can be carried out using software such as Praat (Boersma & Weenink, 2016) and Audacity (Audacity, 2014).

Word-level stress is assessed in a similar manner. A common characteristic of dysarthria and AOS is stress on the wrong syllable in a multisyllabic word. To assess lexical stress, the client should produce multisyllabic words in isolation, sentences, and paragraphs to determine whether breakdown occurs at different levels of task complexity.

Prosodic phrase boundaries are perceptually salient when there is a decrease in loudness, lengthening of the final word, a reset in pitch, and/or a pause. The clinician should listen to the client speaking and reading to perceptually identify phrasal boundaries, followed by visual inspection of the acoustic signal to locate and measure the duration of breaks in speech that mark pauses. Breathing patterns should also be noted to assess if the client takes a breath at every break or pauses in the middle of a word or phrase to breathe.

If it is not easy to identify where stress and prosodic boundaries are occurring in speech, then it is likely that the client is not effectively controlling pitch, loudness, and timing for prosodic expression. The next step is to identify which speech subsystems are contributing to lack of prosodic control. Lack of control in coordination of the respiratory and phonatory subsystems is the most likely cause of prosodic impairment in adults, however, nasality and articulatory error can also be contributing factors. Understanding the contribution of these speech subsystems to prosodic control will be important for guiding treatment goals.

Severity of prosodic impairment can be assessed by determining the number of prosodic features that are disrupted and how much they impede the perceptual salience of phrasal stress and prosodic boundaries. For example, a person with ataxic dysarthria will have a severe prosodic impairment if they are speaking in two-word phrases where each word is lengthened with a large change in pitch and loudness. For this client, it will be very difficult to identify stress and boundaries and to make accurate interpretations of prosodic meaning. Alternatively, a person with flaccid dysarthria will have a mild prosodic impairment if articulatory errors from muscle weakness disrupts the timing of multisyllabic words but phrasing, pitch, and loudness are all perceived as normal.

Expression of emotion during speech should also be included as part of an assessment of prosody. Despite the lack of standardized protocols for assessment of emotional prosody, clinicians may conduct an informal assessment of emotional expression by instructing clients to produce target phrases, sentences, and/or passages while conveying a particular emotion. It is often beneficial to have caregivers or family/friends of the patient provide input on the patient's expression of emotion in the case of acquired disorders as they may be able to provide important information about the patient's baseline expression of emotion during speech.

Generally, tools for the assessment of prosody are far more limited than tools to assess other aspects of speech and language (e.g., vocabulary, syntax). As such, assessment of prosody often relies on clinician ratings and judgement. The Prosody Voice Screening Profile (PVSP; Shriberg et al., 1990) may be used to guide assessment of a speaker's phrasing, rate, stress, loudness, pitch, and voice quality. It uses cut-off scores to rate a speaker's prosody as typical across these areas and provides over 200 samples of speech to compare speaker vocalizations against. While adult samples are included, the majority of the samples are produced by children. There is a clear need for additional assessment tools focused on prosody in adults to promote standardization within the field.

6.2. Treatment

There are few evidence-based interventions for prosodic impairment, and most of these interventions treat prosody indirectly. For example, Lee Silverman Voice Treatment (LSVT) focused on increasing vocal loudness in dysarthria (Sapir et al., 2006). Many studies have measured improvements in pitch and timing in speech as well (Levy et al., 2020). Given limited number of treatment studies on prosodic impairment, linguistically informed observations of the perceptual salience of stress and boundaries are useful to guide treatment.

A scaffolding approach for education and treatment of prosodic control is one recommended approach. The clinician should describe how pitch, loudness, and timing are used in speech to emphasize the stressed word relative to surrounding words. Then, the clinician may demonstrate exaggerated phrasal stress and ask the client to identify the stressed word, with attention to changes in pitch, loudness, and timing to build perceptual skills. The same approach can be used for prosodic phrase boundaries. At the end of this stage, the client should be able to consistently identify phrasal stress and prosodic boundaries in the clinician's speech and describe how pitch, loudness, and timing are being used.

The next step is to use short, simple, structured sentences for the client to practice emphasizing the stressed word and pausing appropriately for prosodic boundaries. Information from the speech evaluation should be applied at this level to determine which speech subsystems are contributing to the prosodic impairment. If poor respiratory control is contributing to bursts in loudness, then the client can practice control of expiration during short sentences to modulate loudness for the stressed word. Additionally, planning out the locations of pauses and breaths during this structured task will assist with producing phrasal boundary. If hypernasality is resulting in poor expiratory control, then options such as a palatal lift should be investigated. Common subsystem impairments contributing to prosodic deficit include weakness and/or incoordination in the respiratory, phonatory, or resonatory systems. Acoustic biofeedback may be beneficial for allowing the client to visualize changes in pitch and loudness for the stressed word. Principles of motor learning for increased volume of practice trials and variable feedback should be applied to improve generalizability and learning. At the end of this level, the client should be able to consistently produce salient phrasal stress and boundary in short, simple sentences.

The scaffolding approach proceeds by increasing the complexity of the speech stimuli until the client is able to start implementing the strategies in spontaneous speech. At each level, the clinician must provide verbal comments and biofeedback. For this approach to succeed, the client must have the cognitive ability to consciously change their speech patterns, which requires significant cognitive load. If cognition is impaired, then it is possible that other programs like LSVT may be more beneficial. Another potential approach would be to us minimal pairs for intonation to treat distinctions between statements vs. questions in English, for example.

Intervention for adults with autism may also draw upon the techniques described above. While a given client's focus (e.g., intonation pattern, stress, rate, loudness) in therapy may vary, intervention focused on prosodic production typically includes an emphasis on expression as well as self-monitoring to promote generalization of skills. Several studies, both within and outside the area of speech and language, support the use of visual aids to promote communication skills in individuals with ASD (Elwell, 2019; Kidder & McDonnell, 2017). In line with this research, visual aids representing a particular prosodic characteristic may be used during intervention. For example, a clinician may underline or ask the patient to underline the stressed portion of a word when targeting word-level stress. Additionally, audio recording the client's production of targets, listening together, and discussing the sample together may provide more concrete feedback than verbal feedback alone. Devoting time to develop a client's awareness of their own prosodic differences and the ability to rate changes in their productions within sessions may be particularly important in supporting generalization of skills.

7. Social Factors: association between prosody and interactional dynamics

Prosody plays an important role in interactive speech, through its dual functions of signaling linguistic structure and pragmatic meaning. To put it simply, *how* something is said impacts the interpretation of the *what* was said (Couper-Kuhlen & Selting, 1996). As such, prosodic deficits have the potential to significantly disrupt overall social interactions and thereby warrant assessment and intervention. Prosody is especially important in conversation for turn management (Bögels & Torreira 2015; Ward, 2019). Within an utterance, a speaker may mark the location of a potential turn transition by slower tempo, rising or falling pitch movements, lower intensity, hypoarticulation, creaky voice, and silent pause or audible exhalation at the end of an utterance. A different set of prosodic patterns, e.g., sustained mid-level pitch, are used to signal that a speaker intends to hold the floor beyond the current utterance. A speaker who flouts these prosodic conventions to signal turn-taking may be seen by their interlocutor as uncooperative, or attempting to dominate the conversation.

Considering the important role of prosody in social interactions, there is a burgeoning interest in prosodic entrainment, a phenomenon by which a speaker converges to the prosodic patterns of their conversation partner. Prosodic entrainment can be seen in converging speech rate, voice quality, pause behavior, and pitch patterns and other features over the course of an interaction. Stronger entrainment reflects mutual positive attitudes of the interlocutors and is positively correlated with perceived social attractiveness, mutual likability, competence and supportiveness (Beňuš, 2014). Entrainment between conversation partners is associated with smoother and more successful interactions, with shorter gaps between turns, fewer interruptions and less overlap, and increased objective ratings of success (Levitan et al., 2012).

Atypical prosody arising from prosodic deficits associated with neurological function may impair the speaker's ability to entrain. Likewise, a speaker whose conversation partner exhibits atypical prosody may have difficulty entraining to the novel or irregular prosodic patterns. Deficits in prosodic entrainment are observed in adolescents and young adults with autism (Patel et al., 2020), and in adults with dysarthria (Borrie et al., 2015). Difficulties in assessment have essentially ruled out a focus on prosodic entrainment in clinical practice, although recent developments in the acoustic measurement and modeling of prosody may stimulate new interest in clinical applications (Borrie et al., 2015; Patel 2020).

8. Conclusion

It is fair to say that prosody is the interface that links linguistic structure, discourse meaning, speaker stance, and social dynamics through the modulation of the melodic, rhythmic and energy elements in speech. This chapter has reviewed the function of prosody in marking juncture between words and phrases, and in conveying pragmatic meaning related to reference and speech acts. Prosody was characterized in phonological representation in terms of the prominences and boundaries of hierarchically organized prosodic structures (words and phrases), and phonetically described in speech articulation and acoustics. Physiological requirements for prosody were reviewed for an understanding of how and why prosody deficits may arise in populations with structural or neurological impairment affecting speech, as illustrated in brief for atypical prosody in dysarthria, apraxia of speech, right hemisphere disorder, and autism spectrum disorder in adults**.** Approaches to clinical assessment and intervention were reviewed, with emphasis on approaches that are informed by linguistic research on prosody. Finally, a case was made for a focus on prosody in the clinical domain based on its importance for successful communication and social interaction, as seen in behaviors related to conversational turn management and entrainment.

References

Ackermann, H. (2008). Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. *Trends in Neurosciences*, *31*(6), 265–272. https://doi.org/10.1016/j.tins.2008.02.011

Audacity, T. (2014). Audacity.

Australasian Society for the Study of Brain Impairment, & Joanette, Y. (2015). *Montreal Protocol for the Evaluation of Communication (MEC): manual; 1 stimulus book; 25 copies of the informant questionnaire; 25 copies of the interviewer questionnaire; 5 copies of the response booklet; 1 memory stick with 2 audio files, informant questionnaire, interviewer questionnaire, response booklet*. ASSBI Resources.

Baltaxe, C. A. M., & Simmons, J. Q. (1985). Prosodic Development in Normal and Autistic Children. In *Communication Problems in Autism* (pp. 95–125). Springer US. https://doi.org/10.1007/978-1-4757-4806-2_7

Baltaxe, C., Simmons, J. Q., & Zee, E. (1984). Intonation patterns in normal, autistic and aphasic children. *Proceedings of the Tenth International Congress of Phonetic Sciences*, 713–718.

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, *46*(3–4), 252–267.

Baron-Cohen, S., & Staunton, R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language*, *14*(42–43), 241–248. https://doi.org/10.1177/014272379401404216

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

Beňuš, Š. (2014). Social Aspects of Entrainment in Spoken Interaction. *Cognitive Computation*, *6*(4), 802–813. https://doi.org/10.1007/s12559-014-9261-4

Boersma, P., & Weenink, D. (2016). Praat software. university of amsterdam. *Im Internet: http://www. fon. hum. uva. nl/praat*.

Bögels, S., & Torreira, F. (2015). Listeners use intonational phrase boundaries to project turn ends in spoken interaction. *Journal of Phonetics*, *52*, 46-57.

Borrie, S. A., Lubold, N., & Pon-Barry, H. (2015). Disordered speech disrupts conversational entrainment: a study of acoustic-prosodic entrainment and communicative success in populations with communication challenges. *Frontiers in Psychology*, *0*, 1187. https://doi.org/10.3389/FPSYG.2015.01187

Brazil, D. (1980). *Discourse intonation and language teaching.* ERIC.

Bunn, J. C., & Mead, D. (1971). Control of ventilation during speech. In *JOURNAL OF APPLIED PHYSIOLOGY* (Vol. 31, Issue 6).

Burns, M. S. (1997). *Burns brief inventory of communication and cognition*. Psychological Corporation.

Byrd, D., & Krivokapić, J. (2021). Cracking Prosody in Articulatory Phonology. *Annual Review of Linguistics*, *7*(1), 31–53. https://doi.org/10.1146/annurev-linguistics-030920-050033

Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis, and perception. *Journal of the Acoustical Society of America*, *90*(5), 2394–2410. https://doi.org/10.1121/1.402044

Cho, T. (2005). Prosodic strengthening and featural enhancement: evidence from acoustic and articulatory realizations of /a,i/ in English. *The Journal of the Acoustical Society of America*, *117*(6), 3867–3878. https://doi.org/10.1121/1.1861893

Cole, J., Hualde, J. I., Smith, C. L., Eager, C., Mahrt, T., & Napoleão de Souza, R. (2019). Sound, structure and meaning: The bases of prominence ratings in English, French and Spanish. *Journal of Phonetics*, *75*, 113–147. https://doi.org/10.1016/j.wocn.2019.05.002

Cole, J., Shattuck-Hufnagel, S., & Mo, Y. (2010). Prosody production in spontaneous speech: Phonological encoding, phonetic variability, and the prosodic signature of individual speakers. *The Journal of the Acoustical Society of America*, *128*(4), 2429–2429. https://doi.org/10.1121/1.3508678

Couper-Kuhlen, E., & Selting, M. (1996). *Prosody in conversation: Interactional studies*. Cambridge University Press.

Cutler, A., & Clifton, C. (1984). The use of prosodic information in word recognition. In *Attention and performance X: Control of language processes* (pp. 183–196). http://pubman.mpdl.mpg.de/pubman/item/escidoc:69901:6/component/escidoc:69902/cc_chapErlbaum.pdf

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research*, *12*(2), 246–269.

de Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech*, *36*(2–3), 197–212. https://doi.org/10.1177/002383099303600305

de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America*, *97*(1), 491–504. https://doi.org/10.1121/1.412275

Duffy, J. R. (2013). *Motor Speech Disorders-E-Book: Substrates, Differential Diagnosis, and Management*. Elsevier Health Sciences.

Elwell, E. (2019). *The Impact of Using Visual Supports to Increase Independence for Students with Autism That Require Very Substantial Support in a Vocational Classroom*.

Fay, W. H. (1969). On the basis of autistic echolalia. *Journal of Communication Disorders*, *2*(1), 38–47. https://doi.org/10.1016/0021-9924(69)90053-7

Fudge, E. (2015). *English word-stress*. Routledge.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*(4), 411–458. https://doi.org/10.1016/0010-0285(83)90014-2

Globerson, E., Amir, N., Kishon-Rabin, L., & Golan, O. (2015). Prosody recognition in adults with high-functioning autism spectrum disorders: From psychoacoustics to cognition. *Autism Research*, *8*(2), 153–163. https://doi.org/10.1002/aur.1432

Gordon, M. (2014). Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels. *Word Stress: Theoretical and Typological Issues*, 83–118.

Gussenhoven, C. (2004). *The phonology of tone and intonation*.

Halper, A.S., Bums, M.S., Chemey, L.R., & Mogil, S.I. (1991). *RIC evaluation of communication problems in right hemisphere dysfunction-Revised* (RICE-2). Gaithersburg, MD: Aspen Publishers.

Hirschberg, J. (2015). *Pragmatics and Prosody* (Y. Huang (ed.); Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199697960.013.28

Hixon, T. J., Goldman, M. D., & Mead, J. (1973). Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. *Journal of Speech and Hearing Research*, *16*(1), 78–115. https://doi.org/10.1044/jshr.1601.78

Hixon, Thomas J, Weismer, G., & Hoit, J. D. (2018). *Preclinical speech science: Anatomy, physiology, acoustics, and perception*. Plural Publishing.

Hoit, J. D., Hixon, T. J., Altman, M. E., & Morgan, W. J. (1989). Speech breathing in women. *Journal of Speech and Hearing Research*, *32*(2), 353–365. https://doi.org/10.1044/jshr.3202.353

Holliday, N. (2021). Prosody and Sociolinguistic Variation in American Englishes. In *Annual Review of Linguistics* (Vol. 7, pp. 55–68). Annual Reviews Inc. https://doi.org/10.1146/annurev-linguistics-031220-093728

Huber, J. E. (2008). Effects of utterance length and vocal loudness on speech breathing in older adults. *Respiratory Physiology and Neurobiology*, *164*(3), 323–330. https://doi.org/10.1016/j.resp.2008.08.007

Huber, J. E., & Darling-White, M. (2017). Longitudinal Changes in Speech Breathing in Older Adults with and without Parkinson's Disease. *Seminars in Speech and Language*, *38*(3), 200–209. https://doi.org/10.1055/s-0037-1602839

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders*, *54*(4), 482–499. https://doi.org/10.1044/jshd.5404.482

Kent, Ray D., Kent, J. F., Duffy, J. R., Thomas, J. E., Weismer, G., & Stuntebeck, S. (2000). Ataxic Dysarthria. *Journal of Speech, Language, and Hearing Research*, *43*(1–5), 1275–1289. https://doi.org/10.1044/jslhr.4305.1275

Kidder, J. E., & McDonnell, A. P. (2017). Visual Aids for Positive Behavior Support of Young Children With Autism Spectrum Disorders. *Young Exceptional Children*, *20*(3), 103–116. https://doi.org/10.1177/1096250615586029

Krivokapić, J. (2014). Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. In *Philosophical Transactions of the Royal Society B: Biological*

Sciences (Vol. 369, Issue 1658). Royal Society of London. https://doi.org/10.1098/rstb.2013.0397
Kroeger, P. (2018). *Analyzing meaning: An introduction to semantics and pragmatics*. Language Science Press.
Kummer, A. W., & Lee, L. (1996). Evaluation and treatment of resonance disorders. In *Language, Speech, and Hearing Services in Schools* (Vol. 27, Issue 3, pp. 271–281). American Speech-Language-Hearing Association. https://doi.org/10.1044/0161-1461.2703.271
Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.
~~Levinson, S. C. (1983). Pragmatics Cambridge University Press. *Cambridge UK*.~~
Levitan, R., Gravano, A., Willson, L., Beňuš, Š., Hirschberg, J., & Nenkova, A. (2012). Acoustic-prosodic entrainment and social behavior. *NAACL HLT 2012 - 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, 11–19. http://www.mturk.com
Levy, E. S., Moya-Galé, G., Chang, Y. H. M., Freeman, K., Forrest, K., Brin, M. F., & Ramig, L. A. (2020). The effects of intensive speech treatment on intelligibility in Parkinson's disease: A randomised controlled trial. *EClinicalMedicine*, *24*, 100429. https://doi.org/10.1016/j.eclinm.2020.100429
Mesibov, G. B. (1992). Treatment Issues with High-Functioning Adolescents and Adults with Autism. In *High-Functioning Individuals with Autism* (pp. 143–155). Springer US. https://doi.org/10.1007/978-1-4899-2456-8_8
Murdoch, B. E., Manning, C. Y., Theodoros, D. G., & Thompson, E. C. (1997). Laryngeal and phonatory dysfunction in Parkinson's disease. *Clinical Linguistics and Phonetics*, *11*(3), 245–266. https://doi.org/10.3109/02699209708985194
Navarro, A. H., & Nebot, A. C. (2014). On the importance of the prosodic component in the expression of linguistic im/politeness. *Journal of Politeness Research*, *10*(1), 5–27. https://doi.org/10.1515/pr-2014-0002
Patel, S. (2020). *Verbal Entrainment and Correlated Neural Mechanisms in Autism Spectrum Disorder and First-degree Relatives.* PhD Thesis, Northwestern University.
Patel, S. P., Kim, J. H., Larson, C. R., & Losh, M. (2019). Mechanisms of voice control related to prosody in autism spectrum disorder and first-degree relatives. *Autism Research*, *12*(8), 1192–1210. https://doi.org/10.1002/aur.2156
Patel, S. P., Kritika Nayar, ·, Gary, ·, Martin, E., Franich, K., Crawford, S., Diehl, J. J., & Losh, M. (2020). An Acoustic Characterization of Prosodic Differences in Autism Spectrum Disorder and First-Degree Relatives. *Journal of Autism and Developmental Disorders*, *50*, 3032–3045. https://doi.org/10.1007/s10803-020-04392-9
Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and Language*, *97*(2), 123–134. https://doi.org/10.1016/j.bandl.2005.08.010
Peppé, S. J. E. (2009). Why is prosody in speech-language pathology so difficult? *International Journal of Speech-Language Pathology*, *11*(4), 258–271. https://doi.org/10.1080/17549500902906339
Pimental, P., & Knight, J. (2000). MIRBI-2: The Mini Inventory of Right Brain Injury. *Austin, Tex: Pro-Ed*.
Platt, L. J., Andrews, G., & Howie, P. M. (1980). Dysarthria of adult cerebral palsy: II. Phonemic analysis of articulation errors. *Journal of Speech and Hearing Research*, *23*(1), 41–55. https://doi.org/10.1044/jshr.2301.41
Pronovost, W., Wakstein, M. P., & Wakstein, D. J. (1966). A longitudinal study of the speech behavior and language comprehension of fourteen children diagnosed atypical or autistic. *Exceptional Children*, *33*(1), 19–26. https://doi.org/10.1177/001440296603300104
Roettger, T. B., & Franke, M. (2019). Evidential Strength of Intonational Cues and Rational Adaptation to (Un-)Reliable Intonation. *Cognitive Science*, *43*(7), e12745. https://doi.org/10.1111/COGS.12745
Russell, N. K., & Stathopoulos, E. (1988). Lung volume changes in children and adults during speech

production. *Journal of Speech and Hearing Research*, *31*(2), 146–155. https://doi.org/10.1044/jshr.3102.146
Sapir, S., Olson Ramig, L., & Fox, C. (2006). The Lee Silverman Voice Treatment ® for voice, speech and other orofacial disorders in patients with Parkinson's disease. *Future Neurol*, *5*, 563–570. https://doi.org/10.2217/14796708.1.5.563
Shattuck-Hufnagel, S., & Turk, A. E. (1996). A Prosody Tutorial for Investigators of Auditory Sentence Processing. *Journal of Psycholinguistic Research*, *25*(2), 193–247. https://doi.org/10.1007/BF01708572
Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research*, *44*(5), 1097–1115. https://doi.org/10.1044/1092-4388(2001/087)
Shriberg, Lawrence D, Kwiatkowski, J., & Rasmussen, C. (1990). *Prosody-voice screening Profile (PVSP): Scoring forms and training materials*. Communication Skill Builders.
Sluijter, a M., & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, *100*(4(1)), 2471–2485. https://doi.org/10.1121/1.417955
Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The Influence of Prosodic Structure on the Resolution of Temporary Syntactic Closure Ambiguities. *Journal of Psycholinguistic Research*, *25*(2).
Spencer, K. A., & Rogers, M. A. (2005). Speech motor programming in hypokinetic and ataxic dysarthria. *Brain and Language*, *94*(3), 347–366. https://doi.org/10.1016/j.bandl.2005.01.008
Tardif, M., Berti, L. C., Marino, V. C. D. C., Pardo, J., & Bressmann, T. (2018). Hypernasal Speech Is Perceived as More Monotonous than Typical Speech. *Folia Phoniatrica et Logopaedica*, *70*(3–4), 183–190. https://doi.org/10.1159/000492385
Titze, I. R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *Citation: The Journal of the Acoustical Society of America*, *85*, 901. https://doi.org/10.1121/1.397562
Van Bourgondien, M. E., & Woods, A. V. (1992). Vocational Possibilities for High-Functioning Adults with Autism. In *High-Functioning Individuals with Autism* (pp. 227–239). Springer US. https://doi.org/10.1007/978-1-4899-2456-8_12
Wang, Y. T., Green, J. R., Nip, I. S. B., Kent, R. D., & Kent, J. F. (2010). Breath group analysis for reading and spontaneous speech in healthy adults. *Folia Phoniatrica et Logopaedica*, *62*(6), 297–302. https://doi.org/10.1159/000316976
Ward, N. G. (2019). *Prosodic patterns in English conversation*. Cambridge University Press.
Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, *19*(6), 713–755. https://doi.org/10.1080/01690960444000070
Westera, M., Goodhue, D., & Gussenhoven, C. (2020). Meanings of tones and tunes. In *The Oxford Handbook of Language Prosody* (pp. 443–453). Oxford University Press Oxford.
Yorkston, K., Beukelman, D., Strand, E., & Bell, K. (1999). Management of Motor Speech Disorders in Children and Adults. PRO-ED. *Inc., Austin, TX*.
Zhang, Z. (2015). Regulation of glottal closure and airflow in a three-dimensional phonation model: Implications for vocal intensity control. *The Journal of the Acoustical Society of America*, *137*(2), 898–910. https://doi.org/10.1121/1.4906272
Zyski, B. J., & Weisiger, B. E. (1987). Identification of dysarthria types based on perceptual analysis. *Journal of Communication Disorders*, *20*(5), 367–378. https://doi.org/10.1016/0021-9924(87)90025-6