Carlos Gussenhoven and Aoju Chen (eds.) *The Oxford handbook of language prosody*. Oxford: Oxford University Press, 2020, lvi+892 pp.
(ISBN 9780198832232)
Reviewed by Gerard O'Grady (Cardiff University)

It is pleasure to review *The Oxford Handbook of Language Prosody* which is an innovative and ground-breaking work. No comparable handbook has been previously published. The editors suggests in the introduction that their enterprise has emerged at this time because of the occurrence of three factors. The first is that Pierrehumbert's (1980) model of prosody is widely understood and enables researchers to work within a common framework. The second is the expansion of databases which means that evidence from a wider range of languages is now available. The third is technology which enables new research paradigms to emerge.

The book is divided into 8 parts of varying length. Part I *Fundamentals of Language Prosody* comprises two chapters. The first chapter, by Cho and Mücke, offers a comprehensive review of the techniques used to measure the articulation of prosody. It convincingly details the benefits and limitations of the various techniques, with a useful tabulated summary of these on pages 26 and 27. The chapter concludes by acknowledging that much remains unclear with regard to the balance between biological propensity and the cultural demands of a language in how it is articulated. This is a motif that repeats throughout the volume. The second chapter, Niebuhr, Reetz, Barnes & Yu, outlines the state of the art in research into F0. Usefully for readers less familiar with the area, the authors provide a concise but detailed review of the main theories of pitch perception. They show that while theories that assume that pitch is perceived depending on the place of stimulation in the inner ear can explain much, such theories fail to account for the perception of pitch in the absence of F0. They also note that theories which assume that pitch perception is dependent on the temporal distance between the firing of neurons are more suitable for the perception of the pitch of pure tones than they are for perceiving just noticeable relative differences, i.e. small rises and falls especially at lower frequencies. We can thus assume that pitch perception is likely to depend on the interaction of place and distance, but they note that this has not yet been fully demonstrated. The chapters ends with an apt caution that F0 is not pitch and that visual representations of F0 as the pitch curve are idealisations and may on occasion mislead.

Part II, *Prosody and Language Structure,* comprises 5 chapters, opening with Hyman and Leben's a comprehensive survey of the form and function of lexical tone systems. Their description goes well beyond the usual crude division into contour and register tones and shows that across languages: (i) the number of tones ranges from 1 to 7, (ii) the tone bearing unit varies from morae to words and phrases, (iii) assimilatory and dissimilatory practices are language dependent, and (iv) in some languages tone is chiefly lexical while in others it functions akin to a derivational morpheme. We are left with the thought that knowing one tone language means familiarity with no more than one tone language. The following chapter by Gordon and Van der Hulst, surveys the related topic of lexical stress, and provides a useful reminder that stress is more than the interaction between loudness, duration and F0. The authors note that across languages duration appears to be the most significant indicator of stress, though in English it is F0 height. While languages have been traditionally divided into those with a stress system that is phonologically predictable and those where stress varies as a function of morphology, the authors show that in reality stress patterns in most languages are a blend of the two. As stress is simultaneously a perceptible quality Gordon and van der Hulst note that researchers tend to hear stresses in the way that their native languages have accustomed them to so doing. This raises the issue as to whether or not we can speak of stress as a universal linguistic feature or set of features, or as something imposed on other languages by speakers of languages with stress. With this in mind I would have welcomed an additional chapter in this handbook surveying stress as a perceptual phenomenon rather than a set of articulatory features.

Chapter 6 by Arvaniti and Fletcher outlines the autosegmental-metrical theory of intonation and is of itself a solid and comprehensible introduction to the world of ToBI. Nonetheless it is also a disappointing chapter in that by focusing too narrowly it does not enable the reader to learn about competing theories of intonation e.g. the British School of tone movement, INTSINT, the IPO tradition and hence readers unfamiliar with the field would be unaware of how ToBI approaches differ and agree with other theories. Nor would they gain an appreciation as to why the levels approach to intonation has historically been preferred in North America but not in the UK see (Crystal 1969). Mention is made of the importance of nuclear tone, a key insight from the British School and how it can be reconciled with ToBI approaches (see Gussenhoven (2004) and the line drawings in Pierrehumbert (1980)), but without further context, the point remains disappointingly opaque. Chapter 7 by

McCarthy is a short and pithy introduction to prosodic morphology or word phonology. The prosodic hierarchy Word^ Foot ^Syllable ^Mora is illustrated and various examples from a range of languages are used to exemplify its workings. To make full sense of the chapter readers will need some familiarity with optimality theory.

Part II is completed by Sandler, Lillo-Martin, Dachkovsky and Müller De Quadors' chapter on Sign Language prosody. The authors compare three sign languages in order to demonstrate that non-manual articulations such as facial expressions function akin to intonation in spoken language. They demonstrate the existence of a "prosodic" hierarchy consisting of utterance ^ intonational phrase ^ intermediate phrase ^ word ^ foot ^ syllable and that prosody functions to signal information structure and illocutionary force. They also demonstrate that non-manual patterns vary across the languages. In conclusion Part II of the handbook has nicely demonstrated the existence of a prosodic hierarchy across languages, but also shown that each language makes its own choices from the available hierarchy.

Part III, *Prosody in Speech Production* comprises 3 chapters. The first, by Barnes, Mixdorff & Niebuhr, examines the phonetic motivation for prosodic realisation by surveying production and perception evidence from a range of languages. The authors illustrate how the interaction of physiological constraints with cultural conventions have the potential to result in differing realisations of the same phonological categories both between and within varieties. Chapter 10, by Van Heuven & Turk explores the phonetic correlates of primary lexical stress and nuclear accent. While they chiefly focus on evidence from English and Dutch, their discussion includes evidence from a range of other languages that duration and intensity are the main acoustic, if not perceptual cues to primary lexical stress and that a change in F0 is the primary cue for identifying nuclear accent – thus supporting Bolinger's (1986) distinction between lexical stress and pitch accents. In contrast, they find no cross-linguistic evidence to support the functional load hypothesis which claims that cues used within language to maintain segmental contrast will be less salient when marking stress. The authors are to be thanked for producing a clear and concise appendix which details how to measure the acoustic correlates of stress using Praat software. The final chapter in this section by White and Malisz surveys the thorny issue of rhythm and timing in spoken language. This chapter illustrates that the current experimental evidence is not supportive of a view that rhythm and timing is based on isochronous units. Instead the authors suggest that rhythm

and timing is likely to be both localised and the result of the interaction between information structure requirements and metrical structure.

Part IV, the heart of the handbook, consists of 18 chapters and surveys prosody across the world. In their introduction, the editors point out that the distribution of languages was chosen in order to reflect the typological diversity which existed prior to the onset of European colonialization, though they also include newer varieties of European languages (see page 8). As would be expected the level of detail varies with much known about the prosodic systems of English and other well-studied European languages and much less known about the indigenous languages of Africa, Australia and South America. The comprehensive chapters are far too rich in detail to review independently, and so I will focus on a number of points pertinent to most if not all chapters. The first point to mention is that for a scholar interested in studying the prosodic systems of under-described languages or language families this handbook represents an indispensable starting point and in conjunction with the relevant chapters in Dryer & Haspelmath (2013) will allow the researcher to get a flavour of the complexity and diversity of individual prosodic systems as well as an idea of where the limits of current knowledge lie.

Reading through the chapters is an enlightening, but in some senses, a disconcerting journey through the complexity and diversity of prosody for a reader used to viewing prosody through the lens of English prosody. It is immediately apparent that the more we learn of non-European languages, the less convincing universals based on European languages are, unless of course the posited universals become contingent and abstract. To illustrate, the chapters show that languages in which lexical tone has a high functional load have limited intonation patterns compared to languages where lexical tone has a lower functional load. Thus, it seems that there is no such thing as an intonationless language, though chapter 12 reports that in some of the highly tonal sub-Saharan languages, intonation is of minimum importance.

Lexical tone where present functions lexically to distinguish minimal pairs such as in Mandarin but in other languages such as most Cushitic ones tone functions grammatically rather than lexically. Prosodic phrasing and prominence similarly seem to occur, though it is reported in Chapter 25, which surveys the extensive Austronesian family, that many of these languages seemingly lack word-based prosody. Metrical feet appear to be common with the majority of surveyed languages favouring trochees rather than iambs; a noticeable exception

however being the languages of North America which, regardless of language family, share a highly marked prevalence for iambic stress patterns. Stress, where it exists, is most commonly signalled by duration. Yet, astonishingly in Pirahã unstressed vowels are reported as having longer duration than stressed ones. Terena is reported to have two means of signalling primary stress: intensity and duration. Stress is fixed in some languages and not in others, and is assigned from the left edge in some languages and the right edge in others. In some languages prominence patterns are more complicated. For instance, Chapter 24 reports that in Tokyo Japanese we find the prominence pattern KA.na.da (Canada), whereas in the Southern Kagoshima dialect we find ka.NA.da. This distinction arises because the former counts prominences in terms of morae and the latter in terms of syllables. We would instinctively expect that languages might divide into those with stress and those with tone but the handbook makes plain that this is an oversimplification with many languages in South America containing both stress and tone.

Intonation, prominence and prosodic phrasing commonly play a role in assigning information focus, in signalling grammatical mood, and in indicating whether an utterance is complete or not. But that the contributions show that nuclear prominence does not necessarily signal 'newness', nor falling intonation a declarative, nor rising intonation an interrogative or incompleteness. For instance chapter 12 reports some of the minor Niger-Congo languages employ intonation to signal grammatical mood, with falling intonation signaling interrogatives (Chapter 12). The same is reported for Mayan languages in chapter 24 and Australian languages in chapter 26. As Austronesian languages of Indonesia apparently lack a system of word prominence, there can be no relation between lexical prominence and 'newness'. Similarly as hinted above in Japanese prominences are not related to discoursal needs. Moreover, in sub-Saharan languages, focus is largely a morphosyntactic feature, while in Georgian it seems to result from the interaction of word order and prosody (chapter 14). In Irish (chapter 20) prosodic prominence and clefting signal information focus. As for Hindi (chapter 21) marked focus is signalled by rise-falling intonation rather than falling intonation with the nuclear accent coterminous with the new constituent.

While the chapters provide evidence for some commonality of prosodic systems within languages and between closely related varieties, it is nonetheless striking how neighbouring languages have mutually influenced and reshaped prosodic systems. Thus, the editorial decision to survey languages by geography rather than by genetic relation has been

more than vindicated. For instance, while Arabic is not a tone language, Moroccan Arabic, as a result of contact with Berber, is (chapter 13). Similarly, in chapter 22 we read that the Siberian language Ket has a tone system which is "strikingly similar" to northern varieties of Chinese. Chapter 14 shows that languages distributed across western and central Asia from four different language families share vowel harmony which is used to signal the existence of a phonological word and also lack a system of contrastive stress. Again, this situation can only have arisen through contact and the spreading of features. However, not all things are affected by contact as the languages demark phrases and signal information focus in their own ways. As a final example of prosodic reshaping arising as a result of contact, chapter 16 reports that Cretan Greek declaratives are intoned with a H+L* nuclear accent rather than the H*+L accent typical of Athenian Greek. This is presumed to be the result of long-term contact between Cretan Greek and Italian. Taken as a whole these chapters remind us that when examining the prosodic systems of a language we need to take account of the language's unique history and its patterns of contact with other languages. For instance, I wonder if the differences in prominence patterns between Tokyo and Kagoshima Japanese are due to the lack of contact between speakers of the two varieties in pre-Meiji times.

Chapter 19 examines geographically diverse varieties of English and finds that inner circle forms (Kachru 1986) tend to share systems and have only minor variations. However, this is not the case for outer circle varieties which tend to have prosodic systems that are heavily influenced by the prosodic systems of the indigenous forms such as Hong Kong English or the Englishes of West Africa, the South Pacific and the Caribbean. Once again it seems that a description needs to take account of contact.

Part V, *Prosody in Communication* consists of five chapters. Chapter 30 by Westera, Goodhue and Gussenhoven examines the meaning of tones (prominences) and tunes (intonation movement) in English by contrasting, and where possible, synthesising the differences between generalist and specialist theories of intonation. The authors wisely situate both sets of theories as complementarities and are clearly correct to argue that theories of linguistic intonational meaning cannot be divorced from pragmatics. Kugler and Calhoun, chapter 31, is a solid review of information structure and its relation to prosody. The authors first clarify that focus/background overlaps but is not equivalent to the relational system of given/new. They show that in most languages where prosody is involved in the signalling of information

structure, that the most common cues are prominences. However, in languages where this is not typologically possible the focus may be signalled by prosodic phrasing or pitch register. The discussion of intonational contours strikes this reader as not totally motivated by the prior discussion; signalling a proposition as part of the common ground is after all very different from signalling the referential givenness of lexical items.

Chapter 32 is an all too short review of prosody and discourse. Hirschberg, Beñuš, Gravano & Levitan survey a range of topics: the relationship between prosody and turn talking, the role of prosody in signalling communicative alignment, how prosody indicates emotion, and whether study of the prosodic signal can assist in locating deception. These are all topics which in and of themselves are well deserving of a chapter of their own as the treatment here is too brief to do more than introduce the topics. Furthermore, this chapter and to a lesser extent the previous one, suffer from not being crystal clear as to how appliable the claims are outside of English and other European languages. In chapter 33 Swerts and Krahmer survey visual prosody across cultures. This chapter is an excellent reminder that the perception of speech is multimodal, as the speech signal is overdetermined as a result of evolving to function in both noisy and dark environments. The second part of the chapter examines how visual prosody helps signal differing emotions. This, as the authors admit, is a contentious issue, as the universality of emotions remains a matter of dispute. The final chapter in the section by Sidtis and Yang surveys pathological prosody and carefully lays out what is known about the neurological underpinnings of various language and speech disorders.

Part VI, *Prosody and Language Processing,* comprises 3 chapters. It is unfortunately likely to be read mostly by specialists, which is a shame as it is a fascinating and rich review of the state of the art in this particular area. Chapter 35 by Lau, Xie, Chandrasekaran and Wong surveys what is known about the cortical and subcortical processing of pitch. The authors nicely illustrate how the field has moved on from discussions of lateralisation, and describe the two main competing theories. They show that prior linguistic and musical experience lead to some restructuring of neural pathways -- a process which is epigenetic. In chapter 36, McQueen and Dilley, survey the state of the art in relation to prosody and spoken word recognition. They show that a division of the speech signal into segmental and suprasegmental elements is not helpful as prosodic cues help demark spoken word boundaries. This view in itself is not new, but what is innovative is their proposal for a

Bayesian Prosody Recognizer which would enable hearers to infer spoken words from different elements of the speech signal. The proposed model seems reasonable, but as the authors note has yet to be tested out on non-European languages. The final chapter in this section by Shattuck-Hufnagel, examines the role of phrasing in speech production planning. It builds upon the seminal work of Levelt (1989) in arguing that prosodic structure is involved in the planning of spoken utterances; words are not the only linguistic elements which a speaker plans in advance. At the end of the chapter she helpfully provides a list of outstanding questions and issues – all of which would make a solid PhD. To her list I would add a further question: To what degree does register influence prosodic planning? In Part VI it would have been nice to have had cross referencing with the survey chapters, perhaps this could be done in a second edition.

Part VII, *Prosody and Language Acquisition,* consists of 7 chapters; the first five focus on first language development with the remaining chapters concerned with the second language development of prosody. Fikkert, Liu and Ota in chapter 38 survey the acquisition of word prosody by which they mean lexical tone, pitch accent and lexical stress. The evidence suggests that children by the end of their second year show a preference for the prosodic patterns of their native language, but that in the case of lexical tone children need exposure to tone patterns in order to maintain their ability to discriminate between tonal patterns. Chapter 39 by Chen, Esteve-Gibert, Prieto and Redford provides an in-depth survey of the development of prosody, revealing that while children may appear to develop mastery of intonation patterns at a young age, adult-like production of appropriate sociopragmatic meaning develops slowly and is dependent on the prosodic system of the individual language. The authors note that until now there has been an almost exclusive focus on the prosodic patterns children produce rather than how they do so. Chapter 40 by Gervain, Christophe and Mazuka surveys the open question of the extent of prosodic bootstrapping in the acquisition of an L1. It is clearly a truism that the initial linguistic stimulus available to the foetus is prosody and that prenatal language input occurs. However, as part IV of this handbook elegantly shows, prosodic systems across the world vary enormously and hence it is not settled as to how a bootstrapping mechanism could effectively facilitate lexical acquisition – though it clearly is an important facilitative heuristic mechanism. The authors, as a result, wisely end their chapter by listing a series of open questions. Chapter 41 by Soderstom and Bortfeld examines the relationship between the use of child directed speech (CDS) and

language (especially vocal) development in children across languages. CDS is, the authors note, a particular register of speech identified by unusual pitch and rhythmic patterns. However, its actual realisation varies across languages and perhaps also adapts to the age and gender of the child. As the authors note, despite 50 years of research into CDS, there is still much to learn about its form and to a lesser extent its function.

Chapter 42 by Paul, Simmons and Mashie, reviews the relatively sparse literature on children with atypical development and their prosodic development. Three of the four disorders reviewed, ASD, DLD (also known as SLI) and hearing loss, though compensated by cochlear implants, are caused by a less efficient processing of the speech signal. Yet, the level of prosodic attainment and perception varies within the syndrome suggesting that while the speech signal carries redundant information, not all individuals are able to benefit from its richness. Much less is known about prosodic impairment in the fourth disorder, cerebral palsy, and what is known is based on evidence from adult production. The evidence presented in this chapter shows that sadly, much remains to be learned and that was has been learned is largely based on English. Much more needs to be understood before we hope to see effective treatment and application of research into this area.

The final three chapters of part VII survey prosodic and second language acquisition. Chapter 43 by Jongman and Tremblay surveys word prosody in second language acquisition and demonstrates the importance of experience. Learners whose L1 contains contrastive stress are more likely to perceive and produce lexical stress patterns in the L2. Their ability to approach native-like performance is reported as depending on whether the phonetic cues triggering stress perception are the same between the two languages. Similarly, prior experience with lexical tone helps an L2 learner acquire tone. The authors report that longer exposure in the prosodic patterns of the L2 results in more native-like phonology thus illustrating that with experience, L2 learners can discern the statistical regularities of prosodic patterning from the speech signal. At the end of the chapter, I was left wondering how learner motivation would impact upon their ability to perceive and produce prosodic patterns in the L2. Unsurprisingly, Trouvain and Braun's survey of L2 sentence prosody similarly finds that transfer from the L1 impacts L2 prosodic perception and production. Though as the authors sagely admit, transfer, owing to prosodic variation between languages, is extremely difficult to operationalise and employ. A disappointment with this chapter was that in the discussion about fluency it did not take into account that L2 learners may have been struggling to

produce communicative content owing to their focus being on the language form. The final chapter in part VII by Chun and Levis surveys the place of prosody in L2 teaching. The authors point out that prosody has been neglected in pronunciation theory, but illustrate a number of methodologies grounded in raising awareness of L2 prosody that have proved useful. Taken together these three chapters illustrate that L2 prosody in contrast with L2 phonemic systems is under-researched and hence poorly understood. One key reason may well be the perceived nebulousness of the speech signal and the absence of a transcription system akin to a broad IPA transcription. It can be but hoped that technological development will prove to be of assistance here.

Part VIII, *Prosody in Technology and the Arts,* examines prosody in technology in chapters 46 and 47 and in the arts in the final two chapters. Chapter 46 by Batliner and Möbius, surveys the history of automatic speech prosody and shows that despite early visibility the role of prosody has become less significant. The authors suggest that this is because it is difficult to identify which prosodic features are functional. They also note the importance of power features – the elements that do most of the heavy lifting, such as pausing in identifying intonation phrase boundaries. However, they also note that power features are not always useful in helping human users mimic native-like prosody. In chapter 47 Rosenberg and Hasegawa-Johnson illustrate the workings of an automatic ToBI system. Both of these chapters, solid though they are, are limited by being restricted to English. Furthermore, it is likely that advances in technology will mean that these chapters will not have as long a shelf life as those in Part IV.

The final two chapters focus on the arts with chapter 48 by Kiparsky representing an expansive survey of stress and metre and its setting in verse. Kiparsky assembles a mass of evidence and his use of Shakespearean examples is fascinating. The chapter though requires some knowledge of both AM theory and Optimality theory, so it may prove to be challenging for some readers. The final chapter focused on the relation between tone and melody in tone languages and was eye opening for this reader. Ladd and Kirby's review of the evidence suggests that the relationship between tone and melody is based on local constraints where the melody mimics the movement between adjacent tones. On the evidence they present, this is more than plausible. Being a very greedy reader, however, I would have liked to have seen supplemental evidence of performed songs and verse in this chapter to back up the text-based evidence.

To conclude, this handbook is a monumental achievement which I imagine will be a well-thumbed resource for scholars and students interested in prosody and especially the prosody of lesser studied languages. The editors are to be highly congratulated for producing such a copious and informative work. To finish with, I have two very minor comments which could perhaps be incorporated in a future edition. The first is that the editors or another produce short introductions to each of the parts. This would have been extremely helpful especially for part IV. Finally figures 19.1 and 19.2 are the wrong way around on page 292.

References

Bolinger, Dwight. 1986. Intonation and Its Parts: Melody in Spoken English. Standord CA; Stanford University Press.

Crystal, David. 1969. *Prosodic Systems and Intonation in English.* Cambridge: Cambridge University Press.

Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info, Accessed on 2021-05-21.)

Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation.* Cambridge: Cambridge University Press.

Kachru, Braj 1986. The power and politics of English: *World Englishes*. 5 2-3, 121-140. **doi.org/10.1111/j.1467-971X.1986.tb00720.x**

Levelt, Willem, J. 1989. *Speaking: From intention to articulation.* Cambridge MA: M.I.T Press.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation.* Unpublished doctoral dissertation. M.I.T.