What word-prosodic typology is missing: Motivating foot structure as an analytical tool for syllable-internal prosodic oppositions

Björn Köhnlein[1] · **Ian S. Cameron[1]**

Abstract

A notoriously contested subarea of phonological typology is word-prosodic typology, which governs suprasegmental structure (such as tone, syllable structure and stress) at the word level. Within word-prosodic typology, it is widely recognized that some languages have so-called stress systems while others have lexical-tone systems. Other languages appear to have intermediate systems, with properties of both stress and lexically contrastive tone. Certain types of such intermediate systems are at the core of ongoing theoretical debates on the nature of word- prosodic systems, viz. language varieties with contrasts between two word tones that are restricted to the main-stressed syllables of a word, a phenomenon that is often descriptively referred to as tonal accent. In this paper, we aim to show that exploring tone-accent systems in detail has the potential to significantly contribute to word-prosodic typology, specifically concerning the foot as a tool for the analysis of syllable-internal prosodic contrasts. The phonology of tonal accent in Franconian (a variety of West Germanic spoken in parts of Belgium, Germany, and the Netherlands) will be the main piece of evidence supporting our claims, with a focus on predictable interactions between segmental structure and accentuation. A central implication of our analysis is that tonal contrasts within syllables can sometimes derive from two types of feet being active in the same prosodic system. We support the Franconian evidence with analogous tone-segment interactions in Estonian and discuss the relevance of our claims in the broader context of word-prosodic typology.

Keywords Word-prosodic typology · Phonological typology · Tonal accent · Segment-tone interactions · Metrical structure · Word stress

✉ I.S. Cameron
cameron.996@buckeyemail.osu.edu

B. Köhnlein
koehnlein.3@osu.edu

[1] Department of Linguistics, The Ohio State University, Columbus, OH, USA

⧖ Springer

1 Introduction

Hyman (2009, 213) suggests that "the central goal of phonological typology is to determine how different languages systematize the phonetic substance available to all languages." Regarding the *how*, a notoriously contested subarea of phonological typology is word-prosodic typology, which governs suprasegmental structure (such as tone, syllable structure and stress) at the word level. Within word-prosodic typology, it is widely recognized that some languages have *stress systems* (English being routinely cited as a prototype), while others have *lexical-tone systems* (Mandarin or Thai being routinely cited as prototypes). Other languages appear to have intermediate systems, with properties of both stress and lexically contrastive tone. Certain types of such intermediate systems are at the core of ongoing theoretical debates on the nature of word-prosodic systems, viz. language varieties with contrasts between two word tones that are restricted to the main-stressed syllables of a word, a phenomenon that is often descriptively referred to as *tonal accent*.

In this paper, we aim to show that exploring tone-accent systems in detail may significantly contribute to the central goal of word-prosodic typology as defined by Hyman, specifically concerning the *foot* as a tool for the analysis of syllable-internal prosodic contrasts. While feet have been well-established metrical constituents in the analysis of word stress and rhythm (Hayes 1995, among many others), we argue that word-prosodic typology has not yet fully embraced their potential regarding the analysis of tone (accent) oppositions. The phonology of tonal accent in Franconian (a variety of West Germanic spoken in parts of Belgium, Germany, and the Netherlands) will be the main source of evidence supporting our claims, with a focus on predictable interactions between segmental structure and accentuation. A central implication of our analysis, which we believe is not yet sufficiently recognized in the literature on word-prosodic typology, is that tonal contrasts within syllables can sometimes derive from two types of feet being active in the same prosodic system (in our case: two types of trochees, a moraic vs. syllabic head). We support the foot-based account of the Franconian facts with analogous tone-segment interactions in Estonian, paying special attention to disyllabic words with long (Q2) initial syllables and overlong (Q3) initial syllables. We propose that the typological parallels between Franconian and Estonian are best accounted for with reference to foot structure.

For a basic illustration of Franconian tonal accent, consider first the minimal pairs in (1), which are taken from Cologne Franconian, a variety with an opposition between two tone accents, so-called *Accent 1* and *Accent 2*. The opposition is restricted to main-stressed syllables of prosodic words; furthermore, all relevant stressed syllables with two sonorant moras (long vowels, diphthongs, short vowels plus coda sonorants; see Sect. 2 for further discussion) have either Accent 1 or Accent 2 (like most dialects, Cologne has no accent contrast in syllables with only one sonorant mora, such as short vowels followed by obstruents). In phrase-medial declaratives, Accent 1 is realized as an early fall, and Accent 2 as a high-level tone with a fall after the accent syllable; the precise realization of the accents varies depending on prosodic (position, level of prominence) and pragmatic (type of utterance) factors (see Sect. 2 for further discussion).

(1)   Three tone-accent minimal pairs from Cologne Franconian (Gussenhoven and Peters 2004); accent class is indicated with superscripts
       a. [daːx^1] 'day.DAT'   [daːx^2] 'day.NOM'
       b. [ʃɪŋ^1] 'certificate.PL' [ʃɪŋ^2] 'certificate.SG'
       c. [luːs^1] 'clever'   [luːs^2] 'louse'

In the current literature on word-prosodic typology, we identify two opposing trends as to how such tone-accent contrasts "should" be analyzed. Throughout this paper, we explicate the issue based on work by Larry Hyman and Harry van der Hulst, which we think represent said positions fairly well, and which in turn provides a solid baseline for discussion. We note, however, that there certainly are other relevant contributions to the issue, such as work by Hualde (2006, 2012), Gordon (2016) and Gussenhoven (2018), to name a few.

Arguably the most widespread view on the analysis of tonal accent and related oppositions, as defended prominently in typological work by Larry Hyman (e.g. 2006, 2009), dictates that tonal differences within stressed syllables must be attributed to tonal information being stored in the lexicon. This analytical claim is derived from the observation that word stress appears to be a property of syllables, *not* of units below the syllable (such as moras). Hyman argues that we should therefore expect stress-related contrasts only *between* syllables, but not *within* them. Under this approach, Cologne Franconian would thus have a lexical-tone contrast with (at least) one accent being stored with tone in the lexicon, as indeed claimed in Gussenhoven and Peters (2004) and Peters (2006). A different take on word-prosodic typology can be found in work by Harry van der Hulst (e.g., 2011). Van der Hulst claims that it is in fact possible to have prominence differences within syllables, represented with diacritic accent marks on the first versus the second mora of heavy syllables (such as the first or the second part of a long vowel), contra Hyman's views. According to van der Hulst, such accentual differences can be realized as high pitch targets early vs. late in a syllable.

The foot as a constituent typically does not feature prominently in relevant debates: Hyman discusses feet but largely focuses on their relation to (syllabic) stress, while van der Hulst's model of word-level prosody does not include foot structure at all. We argue here that illuminating the role of the foot as a constituent helps to better understand the place of tonal accent in word-prosodic typology, which in turn has the potential to significantly contribute to the discipline in general. Based on a case study of segment-accent interactions in Franconian and a comparison to existing foot-based approaches to Estonian ternary quantity (e.g., Prince 1980; Odden 1997; Prillop 2013, 2019), we claim that at least some tone-accent systems can be analyzed most parsimoniously as an opposition between two types of feet. Elaborating on earlier work by Köhnlein (e.g., 2016) and van Oostendorp (2017), we differentiate between uneven trochees (with the first syllable being the foot head, deriving Franconian Accent 1) versus moraic trochees (with the first mora being the foot head, deriving Franconian Accent 2). The existence of uneven trochees has sometimes been contested in the literature (as in Hayes 1995), but relevant structures have also been considered useful in the analysis of various phonological phenomena (e.g., Mellander 2003). A prime example in our context is the so-called *Germanic foot*, a trimoraic foot that can consist of a stressed heavy plus an unstressed light syllable (Dresher and Lahiri

1991), comparable to our syllabic Accent-1 foot. For ease of description, we will refer to uneven trochees as syllabic trochees throughout this paper, as this terminology emphasizes the two different types of foot heads employed in our analysis, moraic versus syllabic.

Rather than simply suggesting a trade-off in analytical tools ("tones out, feet in"), we demonstrate that our metrical, foot-based approach can be motivated on the basis of independent, non-tonal evidence, i.e., additional correlates of accent that we argue are best analyzed with feet, rather than with tone. In doing so, we agree with Hyman (2009, 231), who notes that "VERY reduced tone systems... may also be metrical, but that needs to be demonstrated other than by the intuition that culminativity = accentual" (emphasis in original). While the restriction of one accent per word (cf. *culminativity*) may thus be insufficient to conclude that a system is metrical rather than tonal, the additional correlates we discuss consist of a bundle of phonologically relevant properties that interact with the accent distribution. The empirical focus of this paper is on the role of word-medial onset consonants. Consider the examples in (2), all disyllabic items with stress on the first syllable. This structure is typical of relevant dialects; with very few exceptions, native monomorphemic words are maximally two syllables long and always have (stem-) initial stress.[1]

(2)   Interactions of consonant voicing and tonal accent (examples from Aegidienberg; Müller 1900, 4)
        a.   [i:[1].zən] 'iron' ∼ [ri:[2].s-ən] 'tear-INF'
        b.   [ʃu:[1].vən] 'push-INF' ∼ [ʃu:[2].fəl] 'shovel'

In (2a) and (2b), items with intervocalic voiced consonants receive Accent 1, and those with intervocalic voiceless consonants receive Accent 2. These examples reflect a correlation of accent and the voicing of intervocalic consonants that has been a hallmark of a large group of dialects (so-called *Rule-A* dialects), as uncontroversially recognized in, e.g., Nörrenberg (1884); Bach (1921); Gussenhoven (2000); Schmidt (2002); Köhnlein (2015); Boersma (2017); among many others. Over time, later changes have obscured the picture in some dialects, but others have retained this distribution synchronically. While said correlation is "perfect" for Accent 1 in such dialects (i.e., items with voiced intervocalic obstruent/sonorant *always* receive Accent 1), items with intervocalic voiceless consonants typically receive Accent 2, but some items with intervocalic voiceless obstruents receive Accent 1.[2]

Comparable phenomena are found in Estonian which, like other Finnic languages, exhibits a word-medial morphophonological alternation known as *conso-*

---

[1]In this context, the Associate Editor asks about the status of longer borrowed words or loanwords. These areas are still severely understudied, certainly in the dialects we focus on. Unfortunately, tonal accent appears to be on its way out in large parts of the area, certainly in younger generations. Based on our impressions, older speakers are often rather confused or reluctant if they are asked to pronounce loanwords since they usually do not consider them to be part of their dialect.

[2]A note regarding terminology: voiceless obstruents in German dialects are often referred to as "fortis" (because of their relatively "strong" articulation) and realized with aspiration, while voiced consonants are referred to as "lenis" (because of their relatively "weak" articulation, respectively). Such consonants are often produced as plain consonants, rather than with true voicing. For ease of exposition, we will (only) use the widely known labels voiceless/strong and voiced/weak interchangeably, but we could also have employed fortis and lenis.
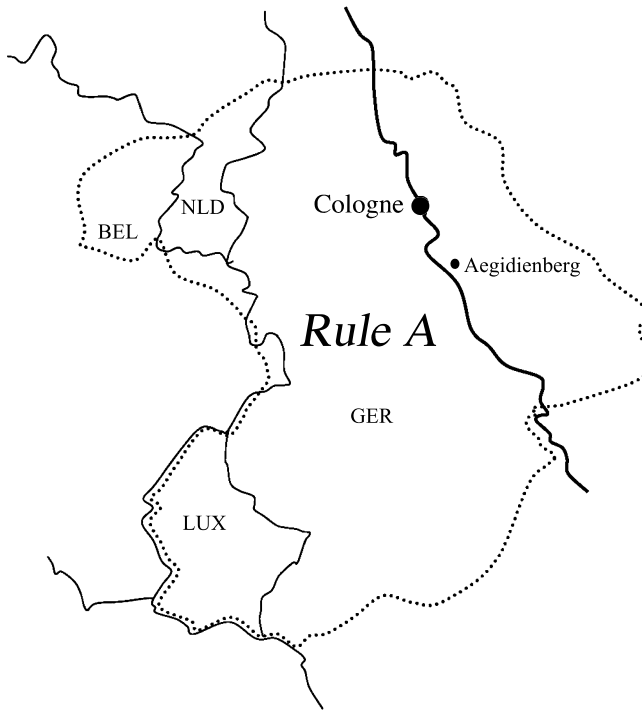
nant gradation. The language also has a ternary length contrast that interacts with gradation; weak grade forms may either be short (Q1) or long (Q2) while strong grade forms are overlong (Q3). For illustration, consider the minimal triplet: *kabi* [kapi] 'hoof.NOM.SG' ∼ *kapi* [kapːi] 'cupboard.GEN.SG' ∼ *kappi* [kappːi] 'cupboard.PTV/ILL.SG.' In certain intonational contexts, the contrast between long and overlong is not only realized by duration, but also by a distinct tonal contrast: Q2 words can exhibit a late fall while Q3 words can exhibit an early fall. As such, the co-occurrence of consonant alternations with distinct accentual patterns offers an interesting parallel with the Franconian facts mentioned above.

While the Estonian patterns have been treated in, e.g., Prince (1980) and Odden (1997), relevant accent-segment interactions in Franconian have not featured in the theoretical discussion to date (at least to the best of our knowledge). As we will argue throughout this paper, however, we believe that they are particularly informative in deciding between available analytical alternatives. After discussing the data in more detail in Sect. 2, we demonstrate in Sect. 3.1 how the patterns can be successfully analyzed with a foot-based approach to the accent opposition. In addition, we briefly discuss the fact that this foot-based account is perfectly compatible with existing metrical analyses of tonal (Sect. 3.2) and durational (Sect. 3.3) correlates of accent (e.g., Köhnlein 2016). In Sect. 4, we then show that our analysis is also compatible with a foot-based approach to ternary quantity in Estonian. In Sect. 5, we compare our approach to alternative analyses with lexical tone (cf. Hyman) or abstract diacritics (cf. van der Hulst). In Sect. 6, we discuss some implications of our analysis for word-prosodic typology and phonological theory. Lastly, we conclude the paper in Sect. 7 by discussing the implications of exploring foot-based analyses of tone-accentual phenomena.

Before we begin our explorations, we note that our foot-based approach to tonal accent aligns with recent literature that analyzes tonal differences between the accents in at least some tone-accent systems as oppositions between two different types of feet, which then leads to different mappings of intonational tones and/or other surface correlates. In addition to the abovementioned Köhnlein (2016) and van Oostendorp (2017) for Franconian as well as Prince (1980), Odden (1997) or Prillop (2013, 2019) for Estonian, this work includes, e.g., Hermans (2012) and Kehrein (2017) for Franconian, Morén-Duolljá (2013), Iosad (2016a, 2016b) for North Germanic, Iosad (2015) and Morrison (2019) for Scottish Gaelic, and Köhnlein (2019a, 2019b) and Köhnlein and Zhu (2019) for Uspanteko. These contributions all share a common analytical thread, i.e., the role of the foot in accentual oppositions, even if they do not always agree on the exact same set of representational tools. We return to this issue at various points throughout the paper, including the conclusion, where we discuss areas for future research.

2 Background and interactions of voicing and accent

In this section, we first provide some general background on Franconian tonal accent and then discuss interactions between consonant voicing and accent, the empirical core of this paper (Sect. 2.1). We end the section by providing the basic tonal melodies and durational correlates of accent based on the Cologne dialect (Sect. 2.2).

Fig. 1 The Franconian and tone accent area (marked by dotted line); adapted from Peters (2006). Rule A covers the geographically largest part of the area (mostly German dialects); other dialect areas are found at the outskirts of the area (particularly in the North East). BEL = Belgium; GER = Germany; LUX = Luxembourg; NLD = The Netherlands

2.1 General background and voicing-accent interactions

Figure 1 shows the Franconian dialect area, spoken in Belgium, Germany, and the Netherlands, and approximately located in the center of the Continental West Germanic area. There is solid evidence that tonal accent has been present in Luxemburgish at some point, but the present-day language appears to have lost the feature altogether (Gilles 2002); Cologne ($\sim$ 1.1 million inhabitants) and Aegidienberg ($\sim$ 7,000 inhabitants) are our examples of dialects which exhibit segment-accent interactions.[3]

When and how tonal accent originated is under debate, with different competing scenarios regarding the emergence of the original system (e.g., Gussenhoven 2000, 2004, 2018; Schmidt 2002; Köhnlein 2013, 2015; Boersma 2017; see Köhnlein 2020 for overview). There is consensus, however, regarding the segmental and metrical contexts under which tonal accent arose: as originally discovered by Nörrenberg (1884) and confirmed many times later, certain sounds and sound combinations in

---

[3]Inhabitant numbers do *not* reflect the number of speakers of tonal accent; the number of speakers in these dialects is unknown but probably much lower than the number of inhabitants, at least in the Cologne case. Tonal accent appears to be on the decline in many dialects, most certainly in younger generations.

older stages of the language accurately predict the accents of present-day dialects.[4] There is a certain amount of distributional variation between different dialect groups, the discussion of which would go beyond the scope of this article; for the purposes of this paper, we focus on the distribution found in Rule A, the geographically most widespread area, which covers most of the dialects spoken on the German side.

First and foremost, as briefly indicated in the Introduction, the accent opposition typically requires two sonorant moras in the stressed syllable, i.e., either a long vowel, a diphthong (both of which we abstractly represent as "VV"), or a short vowel plus a coda sonorant ("VR"). Which accent these units receive depends on several factors: most importantly for our purposes, in Rule A, relevant items *always* receive Accent 1 if a given form was originally disyllabic and had an originally intervocalic voiced consonant, i.e., either a voiced obstruent or a sonorant (which we represent as "D"). Originally disyllabic items typically receive Accent 2 when the intervocalic consonant was voiceless ("T"). This latter generalization applied to stressed syllables with originally long high vowels, diphthongs, and short vowels plus sonorants ("R"); it was overridden, however, by originally long non-high monophthongs, which received Accent 1. The diachronic interactions of consonant quality, vowel quality, and accent are schematized in (3):

(3)　　Interactions of consonant voicing and tone for Cologne and Aegidienberg (= Rule A): diachronic distributional generalizations

　　a.　VVDə, VRDə→ Accent 1

　　b.　VVTə, VRTə→ Accent 2, unless…

　　c.　VV = non-high long monophthong → Accent 1

While these voicing-accent interactions are regular from a historical perspective, later changes have obscured the Rule-A distribution in various dialects: on one hand, a general tendency to lenite obstruents word-medially has neutralized the difference between voiced and voiceless obstruents towards the voiced member in some but, crucially, not all dialects. Moreover, various vocalic changes across dialects (mostly lengthening, but also raising, lowering, diphthongization, monophthongization, centralization) have generally obscured interactions of vowel height and accent, which renders the generalization in (3c) synchronically obsolete. To give an example, the item [ʃlɔː¹fə] 'to sleep' has Accent 1 in Cologne because the form derives from Middle High German *slâfen*, and the long, non-high monophthong triggered Accent 1 by virtue of the diachronic rule in (3c). This rule, however, is probably best regarded as synchronically opaque, one reason being the existence of originally short, but now lengthened vowels with the same quality that receive Accent 2: a case in point is the Cologne item [ʃlɔː²s] 'lock, castle,' which derives from a MHG short vowel (lengthening before a fricative, as per Münch 1904, Sect. 37; cf. MHG *sloz*). It is certainly possible that at some point in the history of the opposition, Accent 1 on old long mid and low vowels was indeed synchronically predictable (Köhnlein 2011, Sect. 7.4 proposes a synchronic analysis of such a stage), but we are not aware of any modern dialect where later vocalic changes have not obscured this regularity. Therefore, we

---

[4]Alternatively, one could use some other suitable reference system, such as West Germanic.

assume that Accent 1 in items with intervocalic voiceless obstruents is synchronically exceptional.

Among the dialects with non-neutralized word-medial obstruents are Cologne and Aegidienberg. Both dialects display the characteristics expected based on the diachronic facts: (i) disyllabic words with word-medial voiced consonants (voiced obstruents, sonorants) always receive Accent 1; and (ii) disyllabic words with word-medial voiceless consonants typically receive Accent 2 but may also receive Accent 1 (if they derive from originally long non-high vowels); recall that monomorphemic words are usually maximally disyllabic and have initial stress. These generalizations are explicitly discussed by Müller (1900, Sect. 3) for Aegidienberg and Münch (1904, Sect. 20, 21) for Cologne. Some examples are given in (4); all examples are transcribed from the original sources into the IPA to increase readability.

(4)   Interactions of consonant voicing and tone for Cologne and Aegidienberg
        a. Cologne (Münch 1904)
           [oː^{1}ɣə] 'eye'   [loː^{2}fə] 'run-INF'
           [lɔː^{1}ɣə] 'site-PL'   [stri^{2}ːfə] 'stripe-PL'
           [freː^{1}zə] 'freeze-INF'   [riː^{2}sə] 'tear-INF'
           [drøː^{1}mə] 'dream-INF'   [ʃliː^{2}sə] 'split-INF'
           [beː^{1}də] 'bid-INF'   [maː^{2}xə] 'do-INF'
        b. Aegidienberg (Müller 1900)
           [iː^{1}zən] 'iron'   [riː^{2}sən] 'tear-INF'
           [ʃuː^{1}vən] 'shove-INF'   [ʃuː^{2}fəl] 'shovel'
           [jaː^{1}ɣən] 'hunt-INF'   [klaː^{2}fən] 'gossip-INF'
           [lyː^{1}nən] 'pay-INF'   [laː^{2}xən] 'laugh-INF'

Regarding expected synchronic exceptions to the generalization, certain items like [ʃlɔː^{1}fə] 'to sleep' have Accent 1 in Cologne: the form derives from Middle High German *slâfen*, and the long, non-high monophthong triggered Accent 1 by virtue of the synchronically opaque Rule in (3c). In addition to these expected exceptions, the relevant sources indicate that some additional cases have arisen over the centuries where we find Accent 2 instead of the expected Accent 1. Such exceptions, however, are largely morphologically conditioned: as observed in Frings (1916, Sect. 24b), they tend to be found in certain verb tenses or adjectival paradigms (specifically in comparatives); therefore, they arguably do not constitute exceptions to automatic phonology. For Aegidienberg, Müller (1900, 12) discusses that comparative forms of adjectives receive Accent 2 even if the word-medial consonant is voiced, an example being [ʃyː^{2}nər] 'more beautiful'; he does not mention any further exceptions. For Cologne, however, Münch (1904) appears to transcribe some apparent non-morphological exceptions.[5] At least in part, however, these appear to be mistranscriptions. For example, [hœənər] 'horns,' which should receive Accent 1 according to Münch's generalizations because of the voiced intervocalic nasal, is once transcribed with Accent 2 (Münch 1904, 36) and once with Accent 1 (Münch 1904, 19); the word [broːdər] 'brother,' which also should receive Accent 1, is once transcribed with Accent 2 (Münch 1904, 143) and three times with Accent 1 (Münch 1904, 55,

---

[5]Some of the relevant items in Münch's grammar have been pointed out to us by Paul Boersma (p.c.).

and twice on page 187). A few words, such as [ʃniːdər] 'tailor,' are transcribed only with Accent 2 (Münch 1904, 42, 53, 90) and thus potentially look more like 'true' exceptions.[6] Such cases are scarce, however, and Münch (1904, 17) himself explicitly states that "*all* long vowels before a voiced consonant and a following unstressed syllable receive the double tone" (= Accent 1; our emphasis and translation).[7] For these reasons, we do not regard this as problematic for our approach—note also that Münch's (incomplete) index contains no less than 1,200 items. In summary, the diachronic distribution of voicing-accent interactions in (3) has been generally preserved in the synchronic Cologne and Aegidienberg systems, which we restate in a general manner in (5); the relevant interactions will be formalized in Sect. 3.

(5)   Interactions of consonant voicing and tonal accent for Cologne and Aegidienberg: synchronic distributional generalizations
   a.   VVDə, VRDə→ Accent 1
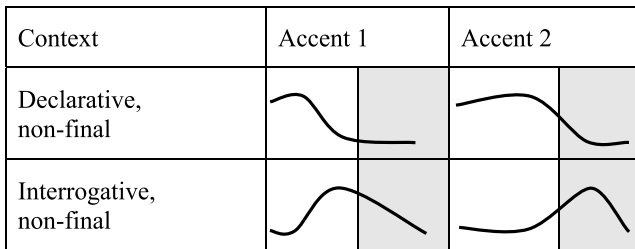   b.   VVTə, VRTə→ Accent 2, unless exceptionally specified for Accent 1

In addition, minimal pairs without a second (schwa) syllable, such as [luːs^1] 'clever' vs. [luːs^2] 'louse,' as shown in (1), will require lexical specification of Accent 1, as their accentuation is not predictable from the segmental content. The analysis of such items will be addressed in Sect. 3.4.

2.2 Tonal and durational correlates

The contrast between the Franconian tone accents is most robustly attested in so-called nuclear syllables but often lost in less prominent positions. In intonational phonology, the notion of a nuclear syllable roughly refers to the syllable that carries maximal prominence in an intonational phrase (under focus). Nuclear syllables can be phrase-final (i.e., located at the right edge of a phrase) or non-final (not located at the right edge of the phrase); they can of course also be phrase-initial, but this position is not relevant for our purposes. For instance, the English intonational phrase *Did you see Justin?* can be used to inquire about a person with that name (underlining indicates focal prominence). Typically, such yes-no questions will be produced with rising, so-called *interrogative* intonation that starts on *Jus-*, the stressed syllable of the name *Justin*; *Jus-* is therefore the non-final nuclear syllable of an interrogative phrase (*-tin* being the phrase-final syllable). In the hypothetical response statement *No, I only saw Ruth!*, *Ruth* is a phrase-final nuclear syllable of the declarative phrase *I only saw Ruth!*, which will typically be marked with falling, so-called *declarative* intonation. All non-nuclear syllables are either in pre-nuclear or post-nuclear position, respectively: in the utterance *I saw Ruth yesterday*, *I saw* would thus be pre-nuclear, *Ruth* nuclear, and *yesterday* post-nuclear.

---

[6]Münch marked accentuation and vowel length with diacritics that are not part of German orthography. For that reason, it is also conceivable that (at least some of) these inconsistencies and possible exceptions may have arisen during typesetting, rather than being attributable to the author (whose descriptions are generally very thorough). Note also that Münch's publisher did not specialize in publishing linguistics books, at least not at the time.

[7]German original: *Den Doppelton haben alle langen Vokale vor stimmhafter Konsonanz mit folgender tonloser Silbe.*

Context	Accent 1	Accent 2
Declarative, non-final		
Interrogative, non-final		

Fig. 2 Idealized tonal contours in Cologne (Rule A), focus, non-final position; accent syllable with nuclear tone-accent contour unshaded, overall post-nuclear contour shaded; relatively wider unshaded box for Accent 2 indicates longer duration of Accent 2

In line with this terminology, Fig. 2 shows idealized tonal realizations of Accent 1 and Accent 2 in Cologne in non-final nuclear syllables of declarative and interrogative intonational phrases. The respective tone-accent syllables with Accent 1 or Accent 2 are unshaded, while the shaded post-nuclear contours represent the general tonal melody after the accent syllables. The precise length/duration of this final stretch varies depending on how much phonological material follows the accented syllable (from one syllable to several post-nuclear items), but the general shape of the pitch contour towards the end of the phrase will be the same. The depictions of the tonal contours are based on acoustic analyses of three minimal pairs from six speakers discussed in Gussenhoven and Peters (2004, Sects. 3, 4) and Peters (2006, Sect. 2), and they closely resemble idealized tonal contours used in earlier work (e.g., Gussenhoven and Peters 2004, 276; Peters 2006, Figs. 11, 12; Riad and Peters 2021, Fig. 18.6). There is prosodically conditioned variation in the realization of the accents across and within dialects (the latter due to pragmatic factors and interactions with boundary tones that mark the edges of intonational phrases; see Köhnlein 2011 for overview), but the contours in Fig. 2 are widely attested across Rule-A dialects and can be considered default realizations of the accents.

Regarding the phonetics of Cologne tonal accent, Gussenhoven and Peters (2004) and Peters (2006) investigated f0, intensity, and duration of accent syllables. While intensity could not be shown to be a consistent correlate (Gussenhoven and Peters 2004, 253), both f0 and duration generally contribute to the realization of the accentual contrast. As can be observed in Fig. 2, the default tonal contours for Accent 1 are an early fall in nuclear declaratives and an early rise in nuclear interrogatives; Accent-2 tonal movements, on the other hand, have a similar shape but occur relatively later: in declaratives, we can observe high level pitch in the nuclear accent syllable and a late fall into the post-tonic contour (shown as grey-shaded in Fig. 2); conversely, we find low level pitch followed by a late rise for Accent-2 nuclear interrogatives. Gussenhoven and Peters (2004, 263) note that "if we abstract away from timing differences, there are no pitch movements which are unique to either accent": accordingly, the tonal difference between Accent 1 and Accent 2 in Cologne lies in the temporal alignment of the pitch contours—relatively early movement corresponds to Accent 1, relatively late movement to Accent 2. The juxtaposition of declarative and interrogative contours also indicates that there is no stable tonal realization of the accents across pragmatic contexts: both Accent-1 and Accent-2 syllables can start

either with high or low pitch, as well as end with high or low pitch. This issue will become particularly important when we compare different approaches to the analysis of predictable interactions between accent and word-medial consonant voicing in Sect. 5.

In addition to f0 differences in nuclear syllables, Cologne shows a consistent durational opposition between Accent 1 (relatively shorter accent syllable) and Accent 2 (relatively longer accent syllable), a property that is widely attested across tone-accent dialects (other correlates, such as differences in vowel quality between the accents, are found in some dialects only). As Gussenhoven and Peters (2004, 257) state, "rhymes with Accent 2 are longer than otherwise identical sonorant rhymes with Accent 1 [. . .] in all positions." Likewise, Peters (2006, Sects. 2.4, 2.6) reports that Accent-2 rhymes are significantly longer than Accent-1 rhymes across all contexts, the difference ranging from 61.9% to 13.8%. In Fig. 2, the general durational difference between the accents is indicated with a relatively wider unshaded box for Accent-2 syllables.[8]

Importantly, Gussenhoven and Peters (2004) and Peters (2006) show that duration is the main correlate of the tone-accent contrast in post-nuclear position.[9] Peters (2006, 19) states that "the f0 difference between Accent 1 and Accent 2 [in post-nuclear position] failed to reach significance."[10] In the absence of a tonal contrast, however, Accent 2 is still significantly longer than Accent 1; this difference is particularly pronounced in the final position of a phrase (Peters 2006, 20). Along those lines, Gussenhoven and Peters (2004, 271) conclude that "the distinguishing phonetic feature in the realisation of the postnuclear contrast is duration." The general durational difference between the accents is shown in (6), where Accent 1 is marked as long and Accent 2 as overlong:

(6) Durational differences between Accent 1 and Accent 2 in the Cologne dialect
     [da:x^1] 'day.DAT' (long)   [da::x^2] 'day.NOM' (overlong)

In summary, we have discussed that the assignment of accent is largely predictable from the voicing of intervocalic consonants in some dialects (examples in (4)), and that Accent 1 and Accent 2 display tonal (shown in Fig. 2) and durational contrasts (examples in (6)). In the next section, we demonstrate how our foot-based approach to accent makes it possible to provide a unified analysis of these phenomena.

3 A foot-based analysis

In this section we first analyze interactions between consonant voicing and tonal accent (Sect. 3.1) and then discuss how our approach derives tonal (Sect. 3.2) and dura-

---

[8]Comparable durational differences are also attested in earlier phonetic work by Heike (1962, 1964); for instance, Heike (1962, 152) characterizes Accent 2 as *lang* 'long' and Accent 1 as *halblang* 'half-long,' the durational difference between the accents ranging from 30% to 55% for different segmental contexts (Heike 1962, Table IV).

[9]No durational data are available for syllables that follow the respective accent syllables.

[10]Gussenhoven and Peters (2004, 273–274) report that there can sometimes be minor pitch differences between the accents in post-nuclear syllables, but they confirm with perception tests that these pitch perturbations cannot be perceived by native speakers.
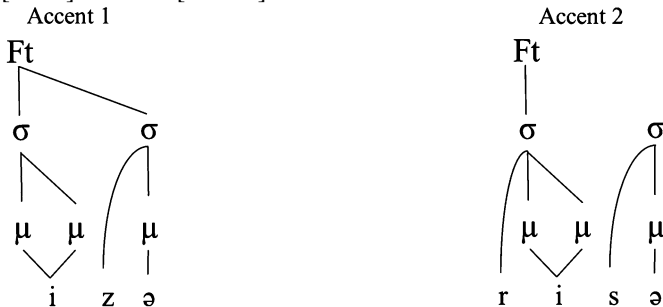
tional correlates (Sect. 3.3) in these disyllabic items. In Sect. 3.4, we discuss how our analysis extends to overtly monosyllabic words. Section 3.4.1 summarizes the key aspects of the analysis.

3.1 How consonant voicing influences foot structure

It is a well-established phonological generalization that domain-medial consonants tend to be "weaker" than consonants at domain edges. For instance, foot-medial onset consonants tend to undergo weakening, a phenomenon commonly referred to as *lenition* (e.g., Honeybone 2012; Katz 2016 for overview). A well-known example is so-called *flapping* in American English, where /t/ and /d/ are neutralized towards a flap in the onset of unstressed syllables that occur in foot-medial position (e.g., *write* vs. *writer*; Davis and Cho 2003). Generally, foot-medial consonants in Germanic languages (and elsewhere) tend to be voiced rather than voiceless (e.g., Holsinger 2000; Smith 2020 for overview). In this subsection, we show that this correlation can successfully be utilized to model voicing-accent interactions in Franconian disyllabic words.

Essentially, we propose that in the Franconian dialects under discussion, word-medial consonant voicing regulates foot structure: on one hand, words with word-medial voiced onsets receive Accent 1, as the foot naturally spans across voiced, "weak" consonants. On the other hand, words with word-medial voiceless onsets typically receive Accent 2 since the voiceless, "strong" consonants in question are preferred at domain edges, and block foot formation across two syllables. Our analysis assumes the default metrical representations in (7). Accent 1, on the left side, shows an item with a word-medial voiced consonant, which results in a disyllabic foot in the dialects in question. Accent 2, on the right side, shows an item with a word-medial voiceless consonant, which blocks the formation of a foot across that consonant, thus preventing a strong consonant from occurring in a foot-medial onset, a position that prefers weaker consonants. The second syllable of the Accent-2 item is unparsed by the foot and links directly to a higher node, presumably the prosodic word (not shown in (7)).

(7)   Influence of consonantal strength on foot structure in Franconian for Cologne [iː^1.zə] 'iron' ∼ [riː^2.s-ə] 'tear-INF'

While such interactions between foot structure and consonant voicing are perfectly in line with typologically based predictions, we note that it has sometimes

been claimed that lenition is always a "top-down" process and should not occur in a "bottom-up" fashion (e.g., Blumenfeld 2006; Rasin 2016). This view thus predicts that certain prosodic positions (such as foot-medial onsets) can influence the realization of segments (= top-down) but that segments are not expected to influence foot structure (= bottom-up). In our analysis, where foot structure is predictable from segmental voicing, these Franconian data would accordingly present counterevidence to this claim; we return to this point in Sect. 6.1.3.

The interactions we postulate can be straightforwardly modeled in any rule- or constraint-based framework which does not exclude the possibility that segmental structure can influence metrical structure. Unless stipulated otherwise, mainstream constraint-based theories like Optimality Theory (Prince and Smolensky 1993) even predict that such interactions should be possible by virtue of constraint reranking. We demonstrate below how relevant patterns can be expressed in OT.

First, the OT formalization requires some constraint against strong consonants in foot-medial position. Such constraints are regularly used to model different lenition patterns, but there are ongoing debates regarding the question how to express cross-linguistically observed patterns in the most parsimonious way (e.g., Smith 2008; Katz 2016). As it is not the goal of our paper to resolve debates on the OT implementation of consonant lenition, we opt for a basic contextual markedness constraint along the lines of Smith (2008) that can easily be replaced with any other constraint from competing frameworks. Furthermore, we assume that the voiceless obstruents in question are specified with a privative feature [SG] (= spread glottis) while voiced obstruents and sonorants are underspecified, following Iverson and Salmon's (2007) specifications for Standard German. This choice is not crucial for the analysis—it would be equally possible to model the interactions with other feature systems, such as binary features. The relevant constraint is defined in (8):

(8)   *FOOT-MEDIAL-[SG] (*FM-[SG]): Assign a violation mark for every foot-medial consonant that is specified for [SG] (spread glottis).

This constraint must outrank PARSE-SYL, a basic OT constraint ensuring that all syllables are parsed by a foot:

(9)   PARSE-SYL: Assign a violation for every syllable that is not parsed by a foot.

These two constraints are sufficient to model predictable interactions of consonantal strength and foot structure. In all tableaux, syllable boundaries will be represented with periods, and foot structure will be indicated with parentheses. As shown in (10) for the item [iː^{1}zə] 'iron,' foot-medial voiced consonants without a [SG]-specification will lead to a disyllabic Accent-1 foot, with Candidate (10a) being the winner (vacuously satisfying *FM-[SG], satisfying PARSE-SYL). The losing Candidate (10b) fatally violates PARSE-SYL.

(10)   Items with word-medial voiced consonants are always footed as a disyllabic Accent-1 trochee

	iːzə	*FM-[SG]	PARSE-SYL
a. →	(iː^1.zə)		
b.	(iː^2).zə		*!

While the tableau in (10) does not provide any ranking arguments, (11) demonstrates for [riː²sə] 'to tear' that *FM-[SG] must outrank PARSE-SYL. This ranking correctly predicts that items with word-medial voiceless consonants (specified for [SG]) will be footed as monosyllabic feet (= Accent 2), followed by an unparsed second syllable, as shown in Candidate (11a); Candidate (11b) with a disyllabic foot fatally violates highly ranked *FM-[SG].

(11)   Items with word-medial strong/voiceless consonants are typically footed as a bimoraic Accent-2 trochee

	riːsə	*FM-[SG]	PARSE-SYL
a.	(riː¹.sə)	*!	
b. →	(riː²).sə		*

As discussed in Sect. 2.1, there is a distinct set of words with etymologically non-high long vowels that always receive Accent 1, i.e., that surface with a disyllabic trochee. To account for such items, we employ a claim in van Oostendorp (2005) and Köhnlein (2011, 2016, 2018): these authors argue based on morphological and phonological evidence that Accent 1 is the marked accent, which we represent here as a lexically stored syllabic trochee (we discuss the issue of how to store unpredictable Accent 1 in underlying representations further in Sect. 3.4 and Sect. 6.2). Following Köhnlein's approach, the heads of such underlying Accent-1 feet are protected by a high-ranked faithfulness constraint HEADMATCH-FT (McCarthy 1995, 2000; Köhnlein 2011, 2016, 2018; Köhnlein and Zhu 2019; Morrison 2019), stated in (12).[11] When HEADMATCH-FT is undominated, it will thus preserve any disyllabic trochee in the input, even if that implies a violation of *FM-[SG].

(12)   HEADMATCH-FT: Assign a violation mark for every element that is a foot head underlyingly but is not a foot head on the surface.
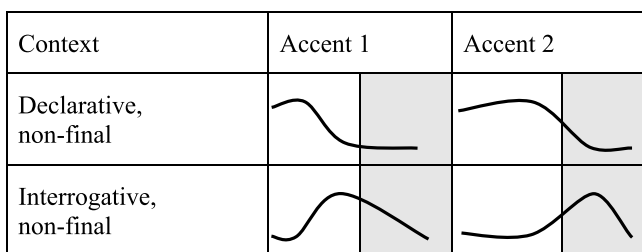
Importantly, forms such as [ʃlɔː¹fə] indicate that the segment-accent interactions in question must be bottom-up, rather than top-down. If this were top-down lenition, we would expect [ʃlɔː¹fə] to either surface with Accent 2 (*[ʃlɔː²fə]) or with Accent 1 and a lenited intervocalic consonant (*[ʃlɔː¹və]). To formally block lenition of /f/ to [v], we introduce another basic OT constraint, a faithfulness constraint preserving the feature [SG]:

(13)   MAX-[SG]: Assign a violation mark for every feature [SG] in the input that is not present in the output.[12]

The relevant interactions are demonstrated in (14) for the item [ʃlɔː¹fə] 'to sleep.' Candidate (14a) wins because it preserves the disyllabic foot that is present in the input (= Accent 1) as well as the voicing quality of the foot-medial consonant, even

---

[11] Formally, we restrict faithfulness to maximal heads, an issue that we do not address here any further but plan to discuss in more detail in future work.

[12] As observed by a reviewer, this constraint does not explicitly rule out the possibility of a delinked feature reassociating to another segment, which could be resolved in various ways (the reviewer suggests replacing MAX with IDENT). As this issue is not at the core of our argument, we will not discuss it any further here.

Context	Accent 1	Accent 2
Declarative, non-final		
Interrogative, non-final		

Fig. 3 Idealized tonal contours in Cologne (Rule A), focus, non-final position; nuclear contour unshaded, overall post-nuclear contour shaded. (Repeated from Fig. 2)

though the candidate violates lower-ranked *FM-[SG]. Candidate (14b), on the other hand, has the preferred 'default foot' that does not span across a voiceless obstruent; however, since its foot head is the first mora but not the first of two successive syllables, it fails to preserve the syllabic foot head in the input, which leads to a fatal violation of HEADMATCH-FT. Lastly, Candidate (14c) preserves the disyllabic foot and satisfies *FM-[SG] but violates undominated MAX-[SG].

(14)   Items with word-medial strong/voiceless consonants *and* lexically stored syllabic trochees surface with Accent 1

	($\int$lɔː^{1}fə) + ('σ.σ)	MAX-[SG]	HEADMATCH-FT	*FM-[SG]	PARSE-SYL
a. →	($\int$lɔː^1.fə)			*	
b.	($\int$lɔː^2).fə		*!		*
c.	($\int$lɔː^1.və)	*!			

To summarize, this subsection has shown how voicing-based footing can be formally accounted for and demonstrated that the interactions in question can be formalized with standard OT mechanisms. In the next subsection, we discuss how foot structure affects the realization of intonational tones.

3.2 How foot structure influences tone

As shown in Sect. 2.2, Accent 1 and Accent 2 in Franconian have contrastive tonal melodies (hence the label *tonal accent*). Following Köhnlein (e.g., 2011, 2016, 2017, 2018) and van Oostendorp (2017), we analyze the relevant tonal oppositions as diverse mappings of the same intonational input tones, which, by virtue of being intonational, are purely postlexical. In other words, the tonal contrast is derived from differences in the association of intonational tones to a metrical opposition between two types of feet, *not* from the presence of tonal information in the lexicon. To illustrate our analysis, we focus on the phrase-medial Cologne nuclear contours introduced in Fig. 2; these contours are repeated for convenience in Fig. 3 (for a detailed analysis and formalization of the overall Cologne tonal system and other relevant dialects in a foot-based approach, consider, e.g., Köhnlein 2011, 83–168).

To reiterate, Accent 1 in Cologne in phrase-medial position is realized as a falling tone in declaratives and as a rising tone in interrogatives; Accent 2 is realized as a high-level tone with a post-tonic fall after the stressed sylla-
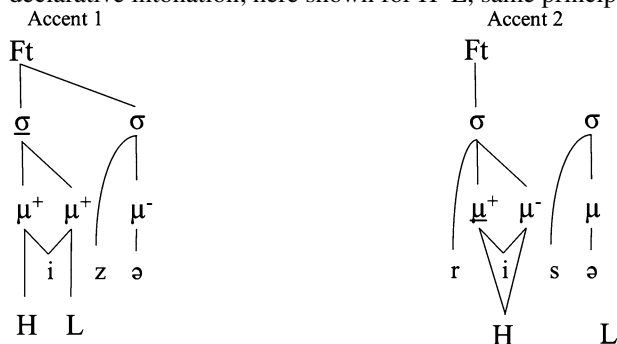
ble in declaratives, and as a low-level tone with a post-tonic rise in interrogatives. Applying basic autosegmental principles, we can assume that contour tones are combinations of two level tones, which means that falls are derived from H and L in the same syllable, and rises from L and H in the same syllable, respectively. Along these lines, we can thus generalize that the Cologne system preferably allows two tones (HL, LH) in Accent-1 syllables but only one tone (H, L) in Accent-2 syllables. In Köhnlein's (e.g., 2011, 2016) approach, this difference follows from the diverse foot structures for Accent 1 and Accent 2.

Before we can discuss how exactly intonational tones map to foot structure, we need to introduce some crucial representational differences between bimoraic, monosyllabic feet and disyllabic feet; this concerns the level of representation at which the respective feet branch. In Köhnlein's approach, foot binarity, the principle responsible for assigning feet a head and a dependent, is determined at the highest level where the foot can be binary (e.g., Morrison 2019 uses the same approach to analyze tonal accent in Scottish Gaelic). Under traditional assumptions in metrical theory, feet can be either binary at the syllabic level or at the moraic level (Hayes 1995). As we have argued in Sect. 3.1, Accent 1 corresponds to a disyllabic foot. Since the foot thus spans two syllables, binarity can be established at the syllable level, with the first syllable being the foot head and the second syllable being the foot dependent. We refer to this type of foot as a *syllabic trochee* (= Accent 1). In (15), left side, headedness in syllabic trochees is shown by underlining the foot head, the first syllable; the second syllable is the foot dependent. Accent 2, on the other hand, corresponds to a monosyllabic, bimoraic foot. The resulting foot is thus binary at the mora level but, crucially, not at the syllable level. Accordingly, the first mora is the foot head and the second mora the foot dependent. We refer to this type of foot as a *moraic trochee* (= Accent 2). In (15), right side, this corresponds to the first, head mora being underlined; the second mora is the foot dependent. Note that all diacritics such as underlining, superscript pluses or superscript minuses are used for purposes of illustration only; they have no representational status in the theory, both in surface and underlying representations.

Crucially, the difference in headedness between the two feet has an impact on the metrical "strength" of the moras in Accent-1 and Accent-2 syllables, respectively. As shown in (15), left side, in syllabic Accent-1 trochees, both moras in the stressed accent syllable are dominated by the foot head, the first syllable. We assume that these moras inherit the strength of the syllabic foot head, which Köhnlein (2011, 2016) refers to as a *head domain*. In other words, Accent-1 syllables contain two moras that are dominated by a foot head, which makes them "strong" at the foot level; in (15), left side, this is indicated with superscript pluses. The mora in the second syllable is dominated by the foot dependent and is therefore metrically "weak" (indicated with a superscript minus). In moraic Accent-2 trochees, on the other hand, the first mora in the accent syllable is the foot head and strong (= superscript plus), while the second mora is the dependent and weak (= superscript minus). Along these lines, Accent-1 syllables thus contain two strong moras, and Accent-2 syllables contain one strong and one weak mora.

(15) Tonal mapping for Accent 1 (left) and Accent 2 (right), phrase-medial declarative intonation; here shown for H*L, same principle with L*H

Accent 1   Accent 2

Ft   Ft

σ̱   σ   σ   σ

μ⁺ μ⁺ μ⁻   μ⁺ μ⁻   μ

i   z   ə   r   i   s   ə

H   L   H   L

With these representations in mind, the Cologne tonal mapping can be derived as follows. The basic principle is that in the Cologne dialect, only strong moras license intonational tones in the accented syllable, and the preferred association is one-to-one. This then affects the mapping of intonational tones, i.e., of the declarative pitch accent H*L and the interrogative pitch accent L*H. Since both moras in a stressed Accent-1 syllable are strong, both tones of the respective intonational pitch accents (H*L, L*H) can be realized in the accented syllable, leading to a fall (H*L) or a rise (L*H). As shown in (15), left side, for declaratives, the first tone goes to the first mora, and the second tone to the second mora. Accent 2, the moraic trochee, only has one strong mora in the accented syllable, which means that it can host only the starred tones of the respective intonational melodies (H*, L*). These starred tones then spread to the second mora (as the tone is already licensed by the first mora, it can spread). The trailing tones (L, H), on the other hand, have to be realized after the accented syllable, the precise details of which depend on the structure of the post-tonic domain (we leave this tone unassociated here, which is not crucial for our argument). In summary, the fact that tonal movements occur earlier in Accent 1 than in Accent 2 follows from our analytical claim that Accent-1 feet license two tones in the accented syllable, and Accent-2 feet license only one tone; crucially, the tonal input for both accents is identical, an intonational pitch accent (either H*L, as in (15), or L*H; the principle remains the same).

While we propose that the tonal differences between Accent 1 and Accent 2 derive from a difference between a syllabic and a trochaic foot, we note that this does not imply that we have to recognize two types of stress—we address this issue in further detail in Sect. 6.1.1. Our analysis does imply, however, that two types of feet can be active in the same language, a claim that we motivate further in Sect. 6.1.2.
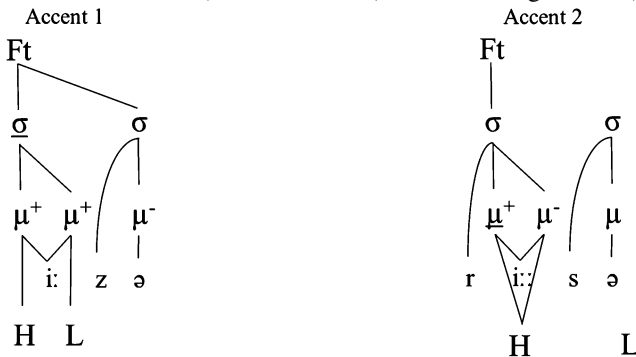
3.3 How foot structure influences duration

In addition to the tonal contrast, Cologne Franconian also displays systematic durational differences between the accents, as do many other Franconian dialects: Accent 2 is substantially longer than Accent 1 (Gussenhoven and Peters 2004; Peters 2006).

As discussed in Sect. 2.2., these durational differences accompany the tonal contrast in nuclear position. In post-nuclear position, however, the durational difference between Accent 1 (long syllable rhyme) and Accent 2 (overlong syllable rhyme) is the relevant correlate of the opposition, rather than pitch.

The use of durational correlates is a well-known property of many metrical (stress) systems, and as such is perfectly compatible with a foot-based analysis of tonal accent. To model this aspect of the tone-accent opposition, we assume that a foot in Cologne Franconian is assigned a certain phonetic duration, which is then distributed across the elements of the foot (based on Köhnlein 2016); as we discuss in more detail in Sect. 4, this analysis is inspired by a foot-based approach to ternary quantity in Estonian (Prince 1980; Odden 1997; Prillop 2013, 2019). In Franconian, the principle can be applied as follows. Accent 1, the (di)syllabic trochee, distributes the duration of the disyllabic foot across the head syllable and the dependent syllable. Accent 2, the monosyllabic, (bi)moraic foot, distributes its duration over only one syllable, the head mora and the dependent mora. Accordingly, Accent 2, where the whole duration of the foot is expressed in the accented syllable, thus has a longer accented syllable than Accent 1, where only some of the duration of the foot is realized in the stressed syllable. This is shown in (16), where the durational differences in the accent syllable are expressed together with the tonal contrast for declaratives.

(16)   Tonal and durational correlates of the opposition between Accent 1 (contour tone, shorter vowel) and Accent 2 (level tone, longer vowel)

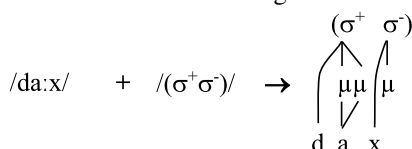## 3.4 The treatment of overtly monosyllabic words

We have argued that the opposition between Accent 1 and Accent 2 is due to a difference in foot structure: Accent 1 is a syllabic trochee spanning two syllables, and Accent 2 is a moraic trochee spanning (only) one syllable. We have argued that this approach not only successfully captures interactions of word-medial consonant voicing and accent (Sect. 3.1) but also successfully models the tonal (Sect. 3.2) and durational (Sect. 3.3) correlates of the opposition. However, not all Franconian items with an accent opposition are overtly disyllabic; recall the examples from (1), which are repeated in (17), now including the durational opposition:

(17)   Three tone-accent minimal pairs from Cologne Franconian, repeated from (1)
   a. [daːx¹] 'day.DAT'   [daːːx²] 'day.NOM'
   b. [ʃɪŋ¹] 'certificate.PL'   [ʃɪŋː²] 'certificate.SG'
   c. [luːs¹] 'clever'   [luːːs²] 'louse'

Following Köhnlein's and van Oostendorp's work, we assume that Accent-1 items with one vowel are followed by a catalectic, empty-headed syllable, i.e., by a syllable with an unpronounced vowel, which is the dependent of a disyllabic foot. In that sense, Accent 1 only appears to be monosyllabic on the surface, but phonologically it is a disyllabic unit. We argue that in such cases, a disyllabic foot is lexically stored, and the empty-headed second syllable is created to ensure its realization as a head. This is comparable to the storage of syllabic trochees in exceptional disyllabic words like [ʃlɔː¹fə] 'sleep' that would otherwise be pronounced with a monosyllabic foot and Accent 2 (as argued for in Sect. 3.1).

There is independent evidence that disyllabic Accent 1 in words with only one vowel is the marked structure: for instance, in all morphologically relevant accent minimal pairs, Accent 1 corresponds to the morphologically more complex category (as observed in van Oostendorp 2005; see also Köhnlein 2016). Regarding the examples in (17), we assume that the syllabic Accent-1 trochee is a morphological exponent in (17a) expressing dative, and in (17b) expressing plural, and stored together with the stem in (17c). We illustrate the derivation of [daːx¹] 'day.DAT' in (18); crucially, because the foot template requires the trochee to be disyllabic (since the initial syllable is an underlying head), the resulting item contains a second syllable with an unpronounced vowel. This makes it possible to faithfully realize the foot template with a syllabic head, which we represent in written text as /($\sigma^+\sigma^-$)/.

(18)   Dative derivation in Cologne: /daːx/ 'day' + /($\sigma^+\sigma^-$)/ 'DAT' → [daːx¹]

Storing foot templates underlyingly, as we propose here, appears to be at odds with the traditional assumption that metrical structure cannot be underlying and/or contrastive—at least not above the level of the mora (e.g., Krämer 2012 for overview). This position derives from the observation that syllabification is normally predictable within a language, and that specifying metrical structure underlyingly appears to predict the possibility of contrastive syllabification. On the other hand, at least within OT, the possibility of storing metrical material underlyingly follows from *Richness of the Base* (e.g., Prince and Smolensky 1993). Furthermore, underlying metrical structure is consistent with the concept of homogeneity of inputs and outputs, i.e., the notion that only objects that can appear in a phonological surface representation can be present underlyingly, and vice versa (Moreton 1999 for discussion). In that sense, we believe that directly storing metrical structure is in principle a viable solution in an approach that employs metrical trees on the surface anyway.

3.4.1 Summary of our analysis

In this section, we have argued that a foot-based model of Franconian tonal accent provides a unified analysis of multiple, independent correlates. Our focus has been on predictable effects of word-medial consonant voicing on accentuation, where we have shown that utilizing well-known interactions of consonantal strength and foot structure provides a straightforward account of the relevant patterns. We have demonstrated that the resulting diverse foot structures are overtly expressed by means of tonal and durational differences and have shown how these correlates can be accounted for in a foot-based approach. In addition to predictable interactions between segmental and metrical structure, we have argued that Accent 1 can also be derived from stored syllabic trochees—this accounts for cases of unpredictable Accent 1 in disyllabic words with word-medial voiceless obstruents, as well as for overtly monosyllabic words that, as we assume, are realized with a catalectic, empty-headed second syllable.
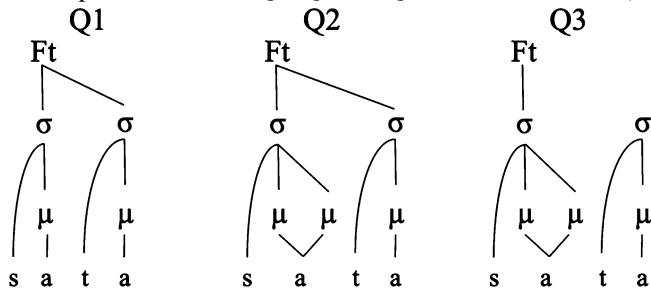
4 Parallels to Estonian quantity

In this section we show that our analysis of segmental-accent interactions in Franconian has close parallels to existing foot-based approaches to Estonian ternary quantity, as in Prince (1980), Odden (1997) and, with certain extensions, Prillop (e.g. 2013, 2019); Prince's original analysis is informed by observations in Hint (1973) that predated foot-based metrical theory but already contained many relevant ingredients for the analysis. An example of the ternary quantity contrast in Estonian is provided in (19), taken from Odden (1997):

(19)   Ternary quantity in Estonian
        a.   [sata] 'hundred' (Q1)
        b.   [saata] 'send-2SG.IMP' (Q2)
        c.   [saːta] 'receive-INF' (Q3)

There is a considerable literature concerning Estonian ternary quantity, including Hayes (1995), Bye (1997), Ehala (2003), Viitso (2003), Pöchtrager (2006), Spahr (2016), or Kuznetsova (2018). Here, we focus on foot-based approaches because this allows for a meaningful comparison to our own analysis of Franconian tonal accent.

The foot-based analysis proposed by Prince, Odden and Prillop is based on the idea that the contrast between Q2 and Q3 can be attributed to two contrasting foot structures, rather than to three degrees of phonemic length: Q3 corresponds to a monosyllabic, bimoraic foot (cf. Accent 2 in our Franconian analysis), Q2 to a disyllabic foot (cf. Accent 1 in our analysis), and Q1 to a disyllabic foot with a monomoraic first syllable—since Franconian dialects typically do not have monomoraic stressed syllables, there is no obvious analogue to the Q1-foot. The corresponding foot structures are given in (20) for the examples in (19); here, we adopt Odden's moraic model of foot structure:

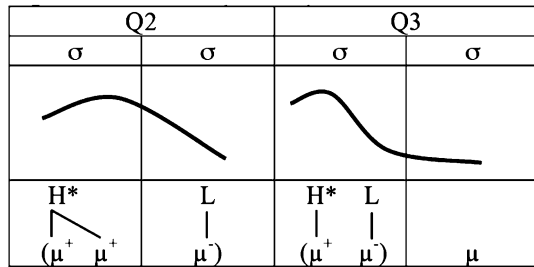(20)   Foot representations of Q1, Q2 and Q3 in a moraic model (Odden 1997)

To account for the durational differences between Q2 and Q3, Prince (1980) proposed that the duration of monosyllabic Q3-feet is expressed in only one syllable, leading to "overlength" in the stressed syllable; in disyllabic Q2-feet, the duration of the foot spreads over two syllables, leading to "normal length" in the stressed syllable—as noted to in Sect. 3.3, we have adopted this principle to account for the durational differences between Cologne Franconian Accent 1 (disyllabic foot, normal length of the stressed syllable) and Accent 2 (monosyllabic foot, overlength of stressed syllable).

In addition to the durational opposition between Q2 and Q3, certain intonational contexts also trigger tonal differences between long and overlong syllables: it has long been observed that the Estonian H*L contour is distributed over two syllables in Q2, while the locus of the H*L contour is entirely in the first syllable for Q3 (e.g., Lehiste 1997; Asu 2004 for overview); we formalize this opposition as shown in Fig. 4. The fact that the durational contrast can coincide with a tonal opposition between an early and a late fall is another obvious parallel between Estonian and Franconian. Interestingly, however, the respective tonal melodies in Estonian and Cologne Franconian are essentially reversed: the longer Q3 in Estonian has a falling tone, while the longer Franconian Accent 2 has a level tone; conversely, shorter Estonian Q2 has a level tone, while shorter Franconian Accent 1 has a falling tone. Therefore, we obviously cannot analyze the tonal mapping in Estonian in the exact same way as in Cologne Franconian.

Notably, however, there are Franconian dialects that display comparable tonal melodies to Estonian, an example being Arzbach Franconian: as shown in Köhnlein's (2011) fieldwork, the disyllabic Accent-1 foot in Arzbach has a high-level tone (plus a post-tonic fall) in declaratives, and the monosyllabic Accent-2 foot has a falling tone, similar to Estonian (and thus the opposite of Cologne). According to Köhnlein (2011, 2016), the Arzbach system can be analyzed by assuming that metrically strong moras, which are linked to the head of a foot, avoid low tone in the dialect, a pattern found in various other tone systems (*HEAD/L; Kenstowicz 1997; de Lacy 2002)—the trailing L of an H*L melody will thus have to be realized on a weak mora. The same principle can account for the Estonian mapping of H*L. As we show in Fig. 4, prohibiting the trailing L from associating to strong moras blocks the low tone from the stressed Q2 syllable: since the syllable node is the foot head, the two moras in the strong syllable of the Q2-foot are both strong, so L must link to the weak mora in the second, dependent syllable. In Q3 syllables, however, the first mora of

Fig. 4 Tonal mapping in Estonian in our foot-based analysis: low tone is avoided on strong moras

Q2		Q3	
σ	σ	σ	σ
H*	L	H* L	
(μ⁺ μ⁺)	μ⁻)	(μ⁺ μ⁻)	μ

the moraic trochee is the foot head and strong, but the second mora is the dependent and weak—the low tone can thus link to the second mora in the stressed syllable, resulting in a falling tone in Q3 syllables.

As briefly mentioned in the introduction, Estonian also features morphophonological consonantal alternations that interact with quantity. This is known as consonant gradation and is characterized by the so-called *strong grade* (with word-medial fortis consonants) versus the *weak grade* (with word-medial lenis consonants). The terms Q1 and Q2 correspond to the weak grade, while Q3 corresponds to the strong grade. For the sake of comparison to the Franconian facts, we will focus on the contrast between Q2 and Q3. Two examples are provided in (21), which shows the inflectional paradigms of the words [kurːp] 'sad' and [teivas] 'pole.'

(21)   Estonian paradigms featuring consonant gradation, taken from Odden (1997, 185)

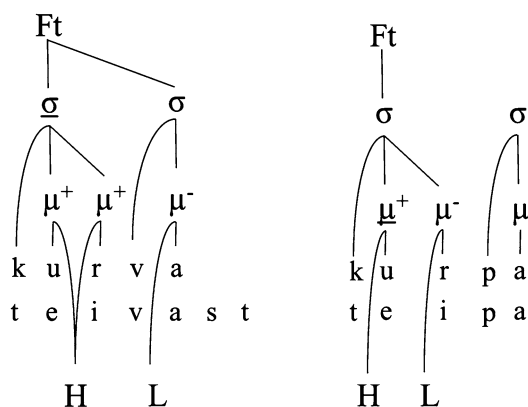|		'sad'	Grade		'pole'	Grade
Sɢ	**Nominative**	kurːp	strong (Q3)		teivas	weak (Q2)
	Genitive	kurva	weak (Q2)		teiːpa	strong (Q3)
	Partitive	kurːpa	strong (Q3)		teivast	weak (Q2)
Pʟ	**Nominative**	kurvat	weak (Q2)		teiːpat	strong (Q3)
	Genitive	kurːpate	strong (Q3)		teivaste	weak (Q2)
	Partitive	kurːpi	strong (Q3)		teiːpait	strong (Q3)

In a foot-based analysis of Estonian prosody, these gradation facts can be accounted for by stating that post-tonic consonants in the middle of the disyllabic Q2 foot are preferably weak, whereas post-tonic consonants after the monosyllabic Q3 foot are preferably strong (see Prince 1980; Odden 1997 for details of the analysis). This is shown in (22)—the gradation facts and their interaction with foot structure are, once again, very similar to what we claim for Franconian. In the disyllabic foot (22), left, two weak grade words are shown with the weak (so-called *lenis*) consonant [v] in foot-medial position and their corresponding tonal melodies mapped to the disyllabic foot along the lines of Fig. 4. In the monosyllabic foot (22), right, two corresponding strong grade words are shown with a monosyllabic foot and its respective tonal melody mapped to the strong and weak moras of the first syllable. The second syllable, whose onset is a strong (so-called *fortis*) consonant [p], remains unparsed by the foot, as in Franconian Accent 2.

(22) Interactions of consonantal strength, foot structure and tonal mapping for Q2 (left) and Q3 (right) in Estonian gradation

Disyllabic Foot (Q2)   Monosyllabic Foot (Q3)

Ft   Ft

σ   σ   σ   σ

μ⁺ μ⁺ μ⁻   μ⁺ μ⁻ μ

k u r v a   k u r p a
t e i v a s t   t e i p a

H   L   H   L

While the prosodic systems of Franconian and Estonian show striking durational, tonal, and segmental similarities, there are also some differences regarding the respective segment-foot interactions: first, Estonian gradation is best treated as "foot-medial weakening" (Odden 1997, 182), i.e., as a top-down process, rather than as a bottom-up process where segmental structure influences footing, as we have proposed for Franconian. Moreover, while the Franconian facts can be shown to be phonologically motivated, this is not as straightforward in the Estonian situation: as discussed in some detail in Odden (1997, Sect. 3.4), certain paradigms do not strictly behave in phonologically regular ways. According to Odden (1997, 185), at least some of these issues can be resolved by specifying foot structure lexically, which again is comparable to our stored syllabic feet deriving Accent 1 in items with word-medial voiceless consonants and in some apparent monosyllables; yet there are also certain alternations in quantity, such as [seipi] 'washer.GEN.SG' with Q2 vs. [seiːpi] 'washer.PART.SG' with Q3 that unexpectedly do not show gradation, even though we do see a change in quantity. Still, the foot-based analysis of the gradation patterns is successful for many paradigms, and as such the similarity to the Franconian facts is remarkable.

Second, it is not in all cases trivial to reduce the Estonian ternary quantity opposition to binary principles, i.e., by attributing the durational difference between Q2 and Q3 solely to different foot structures: there are certain words that appear to contain geminates (typically assumed to be moraic in moraic theory; Davis 2011 for overview) after long or overlong vowels (typically assumed to be at least bimoraic) or after sequences of short vowels plus sonorants. While the precise representational status of these words is not uncontested (Lehiste 1965; Prillop 2013, 2019 for discussion) and the "correct" representation of geminates is still under debate in phonological theory, it is fair to say that such items do not follow straightforwardly in a strictly binary analysis of Estonian quantity (though see Prillop 2013, 2019 for an incorporation of relevant items into the foot-based approach). Yet whatever the status of such items, they do not affect our main points, including their tonal behavior: long Q2 vowels plus a moraic geminate consonant would still contain only strong moras

Table 1 Comparison of Franconian and Estonian foot-based accentual contrasts

	Franconian	Estonian
Foot-based opposition	Disyllabic trochee (Accent 1) vs. monosyllabic trochee (Accent 2)	Disyllabic trochee (Q2) vs. monosyllabic trochee (Q3)
Tonal correlates	Intonational tones map to two feet in diverse ways	Intonational tones map to two feet in diverse ways
Durational correlates	Monosyllabic foot contains overlong accented syllable	Monosyllabic foot contains overlong accented syllable
Segmental correlates	Consonant voicing influences foot structure	Foot structure influences consonant voicing

in the stressed syllable, and hence block the low tone in H*L; overlong Q3- vowels plus a moraic geminate would still associate the low tone to the foot dependent, the second mora, independent of the representation of the geminate—that is, the presence or non-presence of a third mora in some cases does not change our foot-based analysis of the tonal mapping.

In summary, we have aimed to show that our foot-based analysis of Franconian displays several obvious parallels to the foot-based treatment of Estonian ternary quantity in Prince (1980), Odden (1997) and Prillop (2013, 2019), which we believe underscores the typological relevance of our metrical approach to Franconian tonal accent. We briefly summarize the parallels, as established in this section, in Table 1.

In the next two sections, we explicitly return to our overarching themes by discussing the broader relevance of our analysis within word-prosodic typology. In Sect. 5, we evaluate alternative analyses of Franconian tonal accent that would be in line with different analytical trends in word-prosodic typology, as identified at the outset of this paper (lexical tone, diacritic accents). We then discuss the implications of our approach for word-prosodic typology in Sect. 6.

5 Why a foot-based approach is best suited to model the Franconian facts

In previous sections, we have demonstrated how our foot-based analysis of tonal accent allows for a unified analysis of tone, duration and the interaction of word-medial consonant voicing and accent. In the typological overview literature, we have identified two generally available alternatives, an approach that relies on lexical tone (broadly in line with Hyman's approach) and an approach with diacritic markings on moras (broadly in line with van der Hulst's approach). Regarding the lexical-tone approach, there is ample relevant literature available by Gussenhoven and collaborators; here, we focus on Gussenhoven and Peters' (2004) and Peters' (2006) analysis of the Cologne dialect (Sect. 5.1). Analytical suggestions along the lines of van der Hulst's approach can be found in Schmidt (2002) or Boersma (2017, specifically for an earlier, reconstructed stage of Franconian) but detailed analyses are missing; this approach will be discussed in Sect. 5.2.

Fig. 5 Analysis with lexical tone for the Cologne dialect; focus, non-final position. Underlying tonal representation for Accent 2: $/\mu_T\mu/$ ('T' = unspecified tonal node T_{Lex})

Accent 1		Accent 2	
H*	L	H_{Lex}	H*
|	|	|	|
μ	μ	μ	μ
L*	H	L_{Lex}	L*
|	|	|	|
μ	μ	μ	μ

5.1 Franconian tonal accent with lexical tone

Gussenhoven and Peters (2004) (from here, G&P) and Peters (2006) have proposed a tonal analysis of the Cologne facts, and we will use these analyses as a baseline for comparison. G&P argue that the Franconian tone-accent opposition is best analyzed under the assumption that Accent 1 is lexically toneless, and that Accent 2 has a lexical tone: specifically, for Cologne, they claim that Accent 2 has an unspecified tonal node T_{Lex} on the first mora of Accent 2, i.e., a lexical tone that has no underlying value. This tone blocks the assignment of a starred intonational tone (H* in H*L, L* in L*H) to the first mora and forces it to associate with the second mora; the unspecified T_{Lex} then copies the value of this adjacent starred intonational tone. As shown in Fig. 5, the lexical tone thus surfaces as H_{Lex} before H* and as L_{Lex} before L*, respectively. This differentiates Accent 2 with a level tone (H_{Lex} H*, L_{Lex} L*) from Accent 1 with a contour tone (H*L, L*H)—since there is no unspecified lexical tone in Accent 1, no blocking occurs, and both intonational tones can be realized in the accent syllable. In the nuclear syllable of a phrase, the lexical tone thus first acts as a blocking device and then creates a (high or low) level tone in the stressed syllable by copying the value of the following tone, but it does not add an independently observable tonal target.

The analysis successfully captures the empirical tonal facts, though to the best of our knowledge, the proposed interactions would make Franconian the only attested language where intonational pitch accents determine the *quality* of a lexical tone; at least, we have not been able to find any discussion of similar interactions in the literature on, say, tone in tone languages of Asia or Africa. Moreover, recall that in post-nuclear position, the accent opposition in Cologne is realized by means of duration, rather than being accompanied by distinctive tonal differences. G&P (2004, 264) attribute the durational contrast to a process that the authors refer to as "tonal lengthening," which they describe as "phonetically arbitrary." In post-focal position, the unspecified tonal node surfaces solely as extra duration on the syllable rhyme, without providing a distinctive pitch correlate. Once again, we are not aware of other languages where the presence of a (privative) lexical tone can be expressed solely in terms of duration.

Since our approach to voicing-accent interactions is the first synchronic treatment of the phenomenon, we cannot provide an analytical comparison to an existing tonal analysis of the facts. We believe it is fair to note, however, that we find it difficult to identify a straightforward way to account for these synchronic interactions with lexical tone. Of course, it is well-known that high tone on vowels and [-voice] in obstruents correspond to each other in some tone systems, as do low tone and [+voice] (e.g., Bradshaw 2000), so tone-segment interactions are certainly not unexpected. Furthermore, it has also been argued that tone/pitch height can influence voicing of neighboring obstruents, as argued in, e.g., Calabrese and Halle (1998) for Grimm's Law and Verner's Law. In Cologne Franconian, however, Accent 2, the marked accent in G&P's analysis, would be realized as *high tone* in nuclear declaratives (before H*), as *low tone* in nuclear interrogatives (before L*), and as an *unspecified tone* in non-nuclear positions (where it surfaces as extra duration only)—the tonal surface realizations of Accent 2 are thus all over the place and accordingly provide no basis for postulating any consistent tone-segment interaction. Furthermore, the lexical tone is assigned to the first mora, and as such it is unclear how it would affect the voicing quality of a consonant that is separated from said lexical tone by an additional mora and a syllable boundary. Likewise, the tone next to the voiced segment in Accent 1 is either *low* in declaratives (on the second mora after H*), *high* in interrogatives (on the second mora after L*), or *toneless* in non-nuclear position— again, there is no predictable tone associated with the position that is adjacent to the word-medial consonants in questions, which in turn appears to make it impossible to attribute the patterns to surface-transparent tone-segment interactions. This, we believe, is substantially different in our analysis where the relevant interactions can be tied to cross-linguistically well-described differences between foot-medial positions (favoring voiced segments → Accent 1, the disyllabic trochee) versus positions at the boundary of the foot domain (favoring voiceless segments → Accent 2, the bimoraic, monosyllabic trochee).[13]

5.2 Franconian tonal accent with diacritic accents

As indicated in the introduction, there are no detailed analyses of Franconian dialects available in an approach where moras are diacritically marked with accents. Suggestions along these lines have been made in Schmidt (2002) and Boersma (2017), the latter specifically for Old Low Franconian, a predecessor of certain modern Franconian tone-accent dialects. Regarding the tonal facts, present-day dialects like Cologne could certainly be analyzed with moraic accents: Accent 1 would have an accented first mora, and H* and L* could dock to that first mora, leaving room for the respective trailing tones on the second mora. This would lead to an early fall or rise, respectively. Accent 2 on the other hand would have an accented second mora, and the starred tone would dock there, leading to a late fall or rise, respectively. This is shown in (23).

---

[13]We note, however, that Gussenhoven and Peters (2019) have provided a general defense of their approach with lexical tone in comparison to foot-based analyses (though without a discussion of voicing-accent interactions)—also see Köhnlein (2019a, 2019b) for a brief reply.

(23)   Analysis with diacritic marks for the Cologne dialect; μ^* designates a metrically prominent mora that attracts the starred intonational tone

	Accent 1		Accent 2	
Declarative	H* | μ*	L | μ	μ	H* | μ*
Interrogative	L* | μ*	H | μ	μ	L* | μ*

While the Cologne tonal contrast can thus certainly be modeled with diacritics, we are unsure how the durational opposition would be derived in a non-stipulative manner—that is, to derive overlong Accent 2, one would have to limit accentual lengthening to accents on the second mora (though see Spahr 2016 for a multi-layered approach to the diacritic representation of Estonian prosody). More importantly, however, we cannot identify a straightforward way to analyze interactions of consonant voicing and accent. Since diacritic accents do not define domains but are locally specified prominence markers, we do not see a non-stipulative answer to the question why a diacritic accent on the first mora (= Accent 1) should prefer a voiced word-medial consonant after the second, unaccented mora; likewise, we are also unsure why a diacritic accent on the second mora (= Accent 2) should prefer a voiceless intervocalic mora to its right.[14]

In summary, we believe that a foot-based approach is the most promising analytical tool for a comprehensive analysis of Franconian tonal accent, specifically regarding the analysis of predictable interactions of consonant voicing and accentuation. In the next section, we discuss the broader implications of our approach.

6 Implications of our analysis for word-prosodic typology

We have argued here that a foot-based analysis of Franconian tonal accent is a promising analytical approach, with a focus on interactions between segmental structure and accent, which we believe can be expressed most straightforwardly with feet; in addition, we have also accounted for durational and tonal correlates of the opposition and have provided a close typological parallel, viz. Estonian ternary quantity. In this

---

[14] A reviewer wonders if this issue could be resolved by assuming that diacritics could only be used to store the accents underlyingly and would then be translated into foot structure. We discuss the issue of underlying representations in Sect. 6.2 below. Our main argument in this paper, however, regards predictable interactions of segments and accent, however, and as such focuses mostly on surface representations, which do not require underlying storage at all. We indeed believe that stored diacritic accents are not a useful analytical tool for these surface alternations, for the reasons stated in this section.

section, we put our approach in the broader context of word-prosodic typology. In Sect. 6.1, we begin by discussing some implications of our approach, addressing three questions regarding the relation of foot structure to stress (Sect. 6.1.1), the notion of multiple feet in the same prosodic system (Sect. 6.1.2), and the interactions of segmental structure and foot structure we have proposed (Sect. 6.1.3). We then situate our approach in the typological literature, again taking Hyman's and van der Hulst's work as points of reference (Sect. 6.2).

6.1 Some general implications of our approach

6.1.1 Question 1: Does our approach imply that there are two types of stress?

From a formal perspective, our analysis does not imply the existence of two types of stress (even though it might be a useful intuitive notion to think of it that way in a first approximation). It is widely accepted that trochaic feet sometimes count syllables and sometimes count moras (though traditionally considered a language-specific property, on which see Question 2). Yet even in stress systems with mora-counting moraic trochees, the stress-bearing unit is typically identified as the syllable (Hayes 1995, among many others). We take this to mean that our differentiating between syllabic and moraic trochees does *not* necessarily imply that we will have to recognize two types of stress—both moraic trochees (Accent 2) and syllabic trochees (Accent 1) can be assumed to have the syllable as the stress-bearing unit. Rather, we refer to differences in metrical headedness (syllable vs. mora) and domain size (two syllables vs. two moras), which are expected in metrical theory. In our approach, such differences in headedness can indeed lead to prosodic contrasts within stressed syllables (in this paper: tonal and durational contrasts), and the different domains can have effects on post-tonic weak syllables (here: interactions of accent and consonant voicing), but these contrasts are derived from abstract metrical representations and are therefore not directly connected to the notion of stress.

6.1.2 Question 2: Can there be two types of feet in the same prosodic system?

In our approach, the answer is obviously "yes." In traditional, parametric approaches to foot structure (e.g., Hayes 1995), only one type of foot per prosodic system would usually be permitted, based on parameter settings (such as quantity-sensitive vs. quantity-insensitive, iamb vs. trochee, etc.) and indeed, many stress systems can be successfully analyzed by assuming that only one foot is active in the phonology. This does not imply, however, that the notion of "one foot type only" need necessarily be true for all prosodic systems. Empirically, it might not always be easy to find unambiguous evidence for two types of trochaic feet: not all possible candidates are like Estonian or Franconian in demonstrating multiple correlates that (we believe) are perfectly in line with the two-feet assumption and, assuming that the stress-bearing unit is the syllable for moraic and syllabic trochees, speakers would arguably need some additional (phonetic, phonological, distributional) evidence to postulate a difference in footing between two types of words anyway (see Sect. 6.2 for further discussion). Yet relevant cases outside of tonal accent might exist. For instance, Zhu (2023) argues

that the domain for tone sandhi in Suzhou Chinese is either a (bi-)moraic trochee or a (di-)syllabic trochee, depending on the weight of the initial syllable. In addition, there are several languages for which it has been claimed that their metrical systems contain trochees as well as iambs, and where so-called *rhythmic reversals* are caused by certain preferences in the structure of prosodic words (e.g., Prince and Smolensky 1993 for discussion; Bennett and Henderson 2013 for particularly convincing evidence from syncope in Uspanteko). Granted that at least some of these iambic-trochaic analyses are essentially correct, there is thus ample evidence that a given prosodic system can in fact have more than one type of foot. From a theoretical perspective, Optimality Theory even predicts the possibility of multiple foot types per language, by virtue of having rerankable constraints. That is, rankings such as PARSE-SYL ≫ TROCHEE, IAMB in combination with other properties (lexically marked stress, NONFINALITY, etc.) can easily derive oppositions between more than one foot type. We conclude that the possibility of having two types of feet in the same prosodic system can be independently motivated on empirical and theoretical grounds (the latter being the case at least for Optimality Theory).

6.1.3 Question 3: If, in addition to top-down lenition, segmental structure can influence foot structure in a bottom-up fashion, why is there so much cross-linguistic evidence of top-down lenition, but apparently much less evidence of corresponding bottom-up processes?

While, as we have shown in Sect. 3.1, OT-style constraint-based theories generally predict the possibility of having bottom-up interactions of segments and foot structure, it seems clear that lenition is much more robustly attested than the reverse direction. We believe that the reason for this dichotomy may be diachronic, rather than synchronic. In the case of tonal accent, a language learner must have a reason to postulate two types of feet to begin with. That is, the emergence of two foot types will not originate from a conscious decision of learners to set up a more complex prosodic system, but rather from a reinterpretation of existing phonetic tendencies in the speech signal. For Franconian, we know that a complex mix of differences in vowel quality and quantity, word-medial consonant voicing, and changes in syllable number via vowel deletion (= apocope) have contributed to creating the present-day tone accent opposition (Köhnlein 2020 and references therein for overview). In our view, when tonal accent arose as a phonological opposition, the contrast was reinterpreted as an opposition between two types of feet, and in at least some Franconian dialects (i.e., Rule A), this reinterpretation was supported by bottom-up interactions of segments and accent. In many other languages, such specific scenarios required for the emergence of lexically contrastive foot-segment interactions might simply not arise.

That said, broadly comparable phenomena where voicing influences foot structure in a bottom-up manner appear to be attested. That is, there are prosodic systems with so-called onset-sensitive stress, i.e., languages where onsetless syllables avoid stress (e.g., Aranda), where syllables with strong onset consonants attract stress (e.g., Pirahã), or where syllables with weak onset consonants repel stress (Karo)—see Topintzi (2010) for overview. To give an example, in Karo (as described in Gabas 1999,

Sect. 2.5), stress typically occurs on the last syllable of a word, but if that last syllable contains "a voiced stop consonant, /b/, /r/ or /g/, then the stress shifts one syllable to the left" (Gabas 1999, 30; note that Gabas considers /r/ to be a voiced stop phonologically).[15] Blumenfeld (2006) attempts to analyze these effects as top-down lenition, but Topintzi (2010, Sect. 2.2.1.3) shows that such an analysis is empirically problematic. Whatever the "correct" analysis of such onset-sensitive stress systems may be, they demonstrate that bottom-up interactions between consonantal properties and foot structure have been attested for languages other than Franconian.

6.2 Relating our analysis to Hyman's and van der Hulst's approaches to word-prosodic typology

To begin, we obviously share van der Hulst's view that there can be metrically conditioned accentual contrasts within syllables—yet van der Hulst derives such oppositions by assigning accentual prominence to moras (such as the first versus the second mora of a heavy syllable) without involving any constituency. In our approach, however, constituency is key: that is, it crucially relies on the analytical notions of headedness (syllable, mora) and domain (bimoraic, disyllabic) of feet; unlike van der Hulst, we cannot single out a second mora in a stressed syllable as accentually prominent since all feet are built on syllables, including those with moraic heads (see Sect. 3).

As we hope to have shown, the foot as an analytical tool helps to account for Franconian (and Estonian) patterns that suggest evidence of constituency. Since using both feet and diacritic accents in surface representations appears to be undesirable, we argue that surface diacritics can be replaced with foot structure in the analysis of metrically conditioned phonological phenomena. This is in line with other recent metrical work on tonal accent mentioned in the introduction, and one of the core arguments of our paper. As we have alluded to in Sect. 3.4.1, the issue is more complex regarding the question of how unpredictable foot structure is stored in the lexicon, e.g., either as metrical trees or diacritically.[16] Given that lexical representations are further removed from the phonetic realization than surface representations, it is difficult to disentangle these options empirically based on the data we have discussed in this paper, if it is possible at all. What this means for the role of diacritics in prosodic phonology more generally, specifically regarding the structure of lexical representations, is a matter for future research. Answering this question will require a continued theoretical evaluation of competing proposals and their implications, as well as consideration of other types of relevant evidence from various types of prosodic systems. Be this as it may, we believe that our case for foot structure (rather than diacritics only) on the surface is strong for the patterns we have discussed here, and that this argument holds independently of how precisely unpredictable surface foot structure is derived from underlying representations—personally, however, we do see virtue in

---

[15]Nasality and tone also influence stress placement, which we disregard here.

[16]Another possibility that we will not discuss here in detail is storing prominence as a segmental feature [stress] that then gets translated into the maximal head of a metrical tree; see de Lacy (2002) for discussion and a recent proposal along these lines. For our purposes, we can consider this a variant of diacritic marking.

trying to develop a theory where surface and underlying representations are of the same kind.

Moving on to Hyman's approach to word-prosodic typology, arguably the biggest implication of our analysis is that some of Hyman's assumptions are too restrictive—after all, his approach assumes that tonal contrasts within (stressed) syllables must be attributed to lexical tone, which is not compatible with our analysis. We appreciate that a strong appeal of Hyman's diagnostic tool lies in its simplicity: once we observe a syllable-internal tonal opposition in a given prosodic system, the theory predicts that the contrast *must* be due to the presence of lexical tone. In comparison, what we have been pushing for here implies that the analyst and, ultimately, the language learner, do not know immediately whether to attribute a tonal contrast within (bimoraic) syllables to lexical tone or to a metrical opposition. Rather, they have to take other factors into account, such as, in the case of Franconian, whether the tonal opposition is restricted to stressed syllables, whether the number of tonal contrasts is two, whether there are additional correlates accompanying the tone opposition that do not appear to straightforwardly fit into an analysis with lexical tone (e.g., the contrast sometimes being realized solely by durational means), or whether there are interactions of tone and segmental structure that are unexpected under a tonal analysis but fall into place with a foot-based analysis.[17]

While we obviously have no way to conclusively prove our point, we hypothesize that learners might indeed be able to perform such a task, given that they can evidently learn other complex prosodic patterns, such as the diverse phonetic and phonological correlates of stress in Germanic languages (vocalic, segmental, tonal/intonational, distributional) and their realizational variation (correlates depending on phrasal prominence), which they will have to map onto an abstract metrical system (whatever its precise structure). As we have shown throughout this paper, Franconian prosodic systems do in fact use many of the correlates we find in prototypical stress systems, and as such they do not appear to be *that* different from the stress systems of neighboring non-accent dialects and the respective standard languages.

The close similarities between the prosodic systems of accent and non-accent varieties are reflected in, e.g., Ramachers (2018), who finds in a set of experiments that speakers of Dutch (infants and adults) perform remarkably well on the perception of the Franconian tone accent contrast. For instance, Dutch-speaking adults performed significantly above chance in an AXB task that aimed to distinguish Accent 1 and Accent 2, even though such lexical contrasts are not part of their native prosodic system. Ramachers discusses several possibilities for why Dutch listeners performed so well, a likely option being that "Dutch listeners have drawn upon their knowledge of native cues to word stress during their perception of Limburgian tones" (Ramachers 2018, 144)—this interpretation is perfectly compatible with our views and supports the idea that West Germanic word stress and Franconian tonal accent are closely related phenomena. That is, along the lines of our analysis, speakers of Franconian have

---

[17]Conversely, for the existing Cologne analysis with lexical tone in Gussenhoven and Peters (2004), the learner has to infer that the lexical tone does not have a stable realization since it always assimilates to a following intonational tone, that it thus cannot be observed as an independent target, that it can sometimes be realized only by means of duration, and that there are non-transparent interactions with consonant quality, all of which are properties without obvious typological parallels.

prosodic systems that would appear to be largely similar to those of related varieties that lack the lexical distinction between Accent 1 and 2, just that they have more "fine-grained" metrical contrasts which include phonologically relevant oppositions based on the domain of trochaic feet (syllabic or moraic).

In line with these observations, we note that the aspects of Franconian tonal accent discussed in this paper do indeed meet several of the criteria that Hyman (2009, 217) identifies as properties of prototypical stress systems (here *italicized*), viz. *lengthening of stressed syllables* (durational differences between Accent 1 and Accent 2), *non-prominence effects in unstressed syllables* (voiced post-tonic consonants always corresponding to Accent 1, voiceless consonants commonly corresponding to Accent 2), *tonal oppositions being greater on stressed syllables* (no tonal accent on unstressed syllables, reduction of the tonal opposition in non-nuclear syllables), and the *attraction of intonational tones to stressed syllables* (intonational tones being an integral part of the realization of the accents, independent of any specific analysis). For these reasons, we believe that our analysis of Franconian is more easily compatible with Hyman's approach than one might intuitively think. Just to reiterate, the only substantial modification we suggest is that syllable-internal contrasts can sometimes be derived from two types of feet, rather than from lexical tone; as indicated in Sect. 6.1.2, however, this does not necessarily imply that stress itself can affect units below the syllable level, which again is in line with Hyman's typology. Notably, Hyman (2009, 227) himself poses the question "whether metrical structure = stress (as I have assumed), or whether stress = metrical structure + something additional." While Hyman's statement refers not to tonal contrasts within syllables but to the observation that certain languages appear to have foot structure without showing overt correlates of stress, the question as such still indicates that his approach leaves room for a deeper exploration of the role of feet in word-prosodic typology; this in turn is precisely what we have been trying to do throughout this paper.

7 Conclusion: Where to go from here

We have argued throughout this paper that at least some binary tonal oppositions within stressed syllables are best analyzed with reference to constituency, specifically two types of feet being active in the same prosodic system. As such, one typological implication of our approach would be that relevant binary contrasts can be analyzed with reference to either foot structure (syllabic vs. moraic foot) or lexical tone, an analytical decision that may be based on the evidence available to the learner. We have suggested, in line with Hyman's work, that the use of diacritics in surface representations may therefore be obsolete, an argument that can possibly be extended to underlying representations (cf. the discussion in Sect. 6.2), even though this latter issue needs to be evaluated against a wider set of cross-linguistic empirical data than we have been able to consider in this paper.

Assuming that the foot-based approach to at least some accentual word-prosodic oppositions is indeed worth pursuing, it must be noted that this line of research is still in its early days. First and foremost, what is needed are more empirical and theoretical studies of potentially relevant contrasts that should ideally focus not only

on tonal oppositions but explicitly consider distributional restrictions and other phonetic and/or phonological phenomena related to the accentual opposition in question where available. To give but one set of relevant examples, recent metrical analyses of Scottish Gaelic tonal accent (Iosad 2015; Morrison 2019) have deepened our understanding of the prosodic system of the language by explicitly placing tonal accent in the context of various connected segmental and suprasegmental phenomena found in the language. We believe that such detailed studies have the potential to significantly advance our understanding of tonal accent, and as such can also contribute to forming a coherent theory of word-level prosodic structure. For instance, within foot- and mora-based metrical theory, there still are competing approaches regarding the question how exactly relevant representations should be structured. While it has traditionally been assumed that moraic and syllabic feet are both built on syllables, Kager (1993) and subsequent related work (e.g., Kager and Martínez-Paricio 2019) postulate that moraic trochees are directly built on moras, partially based on the analysis of accentual systems; furthermore, there still are debates regarding the question whether foot structure might or might not be recursive (Morén-Duolljá 2013 and Iosad 2016a for examples of using recursive feet to analyze accentual oppositions in North Germanic). We think that down the line, tonal accent as one source of evidence of metrical organization can help to evaluate competing theories of word-prosodic organization.

Along these lines, whatever the "correct" approach to tonal accent will turn out to be, investigating tonal accent and related prosodic oppositions from multiple angles—be it foot structure, lexical tone, accent diacritics, or other tools—will certainly lead to the discovery of novel empirical generalizations and inspire scientific debate, which in turn can only be beneficial for our understanding of "how different languages systematize the phonetic substance available to all languages" (Hyman 2009, 213).

Declarations

References

Asu Eva Liina. 2004. The phonetics and phonology of Estonian intonation. PhD diss., University of Cambridge.

Bach, Adolf. 1921. Die Schärfung in der moselfränkischen Mundart von Arzbach (Unterwesterwaldkreis). *Beiträge zur Geschichte der Deutschen Sprache und Literatur* 45: 266–290.

Bennett, Ryan, and Robert Henderson. 2013. Accent in Uspanteko. *Natural Language and Linguistic Theory* 31(3): 589–645.

Blumenfeld, Lev. 2006. Constraints on phonological interactions. PhD diss., Stanford University.

Boersma, Paul. 2017. The history of the Franconian tone contrast. In *Segmental structure and tone*, eds. Wolfgang Kehrein, Björn Köhnlein, Paul Boersma, and Marc Oostendorp, 27–98. Berlin: de Gruyter.

Bradshaw, Mary. 2000. A crosslinguistic study of consonant-tone interaction. PhD diss., Ohio State University.

Bye, Patrik. 1997. A generative perspective on 'overlength' in Estonian and Saami. In *Estonian prosody: Papers from a symposium*, eds. Ilse Lehiste and Jaap Ross, 36–70. Tallinn: Institute of Estonian language.

Calabrese, Andrea, and Morris Halle. 1998. Grimm's and Verner's laws: A new perspective. In *Mír curad: Studies in honor of Calvert Watkins*, eds. Jay Jasanoff, H. Craig Melchert, and Lisi Oliver, 47–62. Innsbruck: Institut für Sprachwissenschaft, University of Innsbruck.

Davis, Stuart. 2011. Geminates. In *The Blackwell companion to phonology* eds. Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice Vol. 2, 873–897. Oxford: Wiley-Blackwell.

Davis, Stuart, and Mi-Hui Cho. 2003. The distribution of aspirated stops and /h/ in American English and Korean: An alignment approach with typological implications. *Linguistics* 41: 607–652.

de Lacy, Paul. 2002. The interaction of tone and stress in optimality theory. *Phonology* 19: 1–32.

Dresher, B.E., and A. Lahiri. 1991. The germanic foot: Metrical coherence in old english. *Linguistic Inquiry* 22(2): 251–286.

Ehala, Martin. 2003. Estonian quantity: Implications for moraic theory. In *Generative approaches to Finnic and Saami linguistics*, eds. Diane Nelson and Satu Manninen, 51–80. Stanford: CSLI.

Frings, Theodor. 1916. *Die Rheinische Accentuierung: Vorstudie zu einer Grammatik der Rheinischen Mundarten*, Marburg.

Gabas, Nilson Jr. 1999. A grammar of Karo, Tupí (Brazil). PhD diss., University of California, Santa Barbara.

Gilles, Peter. 2002. Einflüsse der Rheinischen Akzentuierung auf die segmentelle Ebene. Evidenz aus dem Luxemburgischen. In *Silbenschnitt und Tonakzente*, eds. Peter Auer, Peter Gilles, and Helmut Spiekermann. 265–282. Tübingen: Niemeyer.

Gordon, Matthew K. 2016. *Phonological typology*, Vol. 1. London: Oxford University Press.

Gussenhoven, Carlos. 2000. On the origin and development of the central Franconian tone contrast. In *Analogy, levelling, markedness: Principles of change in phonology and morphology*, ed. Aditi Lahiri, 215–260. Berlin: Mouton de Gruyter.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos. 2018. Prosodic typology meets phonological representations. In *Phonological typology*, eds. Larry M. Hyman and Frans Plank, 389–418. Berlin and Boston: De Gruyter Mouton.

Gussenhoven, Carlos, and Jörg Peters. 2004. A tonal analysis of Cologne Schärfung. *Phonology* 21(2): 251–285.

Gussenhoven, Carlos, and Jörg Peters. 2019. Franconian tones fare better as tones than as feet: A reply to Köhnlein (2016). *Phonology* 36(3): 497–530.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Heike, Georg. 1962. Suprasegmentale Merkmale der Stadtkölner Mundart: ein Beitrag zur 'Rheinischen Schärfung'. *Phonetica* 8: 147–165.

Heike, Georg. 1964. *Zur Phonologie der Stadtkölner Mundart: eine experimentelle Untersuchung der akustischen Unterscheidungsmerkmale*. Marburg: Elwert.

Hermans, Ben. 2012. The phonological representation of the Limburgian tonal accents. In *Phonological explorations: Empirical, theoretical and diachronic issues*, eds. Bert Botma and Roland Noske, 227–244. Berlin: Mouton de Gruyter.

Hint, M. 1973. *Eesti Keele Sonafonoloogia I*. Tallinn: Eesti NSV Teaduste Akadeemia, Keele ja Kirjanduse Instituut.

Holsinger, Dave. 2000. Lenition in Germanic: Prosodic templates in sound change. PhD diss., University of Wisconsin-Madison.

Honeybone, Patrick. 2012. Lenition in English. In *Handbook on the history of English: Rethinking approaches to the history of English*, eds. Terttu Nevalainen and Elizabeth Closs Traugott, 773–787. Oxford: Oxford University Press.

Hualde, José Ignacio. 2006. Remarks on word-prosodic typology. In *Annual meeting of the Berkeley linguistics society*, Vol. 32, 157–174.

Hualde, José Ignacio. 2012. Two Basque accentual systems and word-prosodic typology. *Lingua* 122(13): 1335–1351.

Hyman, Larry M. 2006. Word-prosodic typology. *Phonology* 23(2): 225–257.

Hyman, Larry M. 2009. How (not) to do phonological typology: The case of pitch-accent. *Language Sciences* 31(2): 213–238.

Iosad, Pavel. 2015. Pitch accent and prosodic structure in Scottish gaelic: Reassessing the role of contact. In *New trends in nordic and general linguistics*, eds. Martin Hilpert, Janet Duke, Christine Mertzlufft, Jan-Ola Östman, and Michael Rießler, 28–54. Berlin: Mouton de Gruyter.

Iosad, Pavel. 2016a. Prosodic structure and suprasegmental features: Short-vowel stød in Danish. *Journal of Comparative Germanic Linguistics* 19(3): 221–268.

Iosad, Pavel. 2016b. Tonal stability and tonogenesis in North germanic. In *Beyond borealism: New perspectives on the North*, eds. Ian Giles, Laura Chapot, Chris Cooijmans, Ryan Foster and Barbara Tesio, 80–98. London: Norvik Press.

Iverson, Gregory K., and Joseph Salmons. 2007. Domains and directionality in the evolution of German final fortition. *Phonology* 24: 1–25.

Kager, René. 1993. Alternatives to the iambic-trochaic law. *Natural Language and Linguistic Theory* 11(3): 381–432.

Kager, René, and Violeta Martínez-Paricio. 2019. Mora and syllable accentuation. In *The study of word stress and accent: Theories, methods and data*, eds. Rob Goedemans, Jeffrey Heinz, and Harry Van der Hulst, 147–186. Cambridge: Cambridge University Press.

Katz, Jonah. 2016. Lenition, perception and neutralisation. *Phonology* 33(1): 43–85.

Kehrein, Wolfgang. 2017. There's no tone in Cologne: Against tone segment interactions in Franconian. In *Segmental structure and tone*, eds. Wolfgang Kehrein, Björn Köhnlein, Paul Boersma, and Marc van Oostendorp, 147–195. Berlin: de Gruyter.

Kenstowicz, Michael. 1997. Quality-sensitive stress. *Rivista Di Linguistica* 9: 157–188.

Köhnlein, Björn. 2011. *Rule reversal revisited: Synchrony and diachrony of tone and prosodic structure in the Franconian dialect of Arzbach*. Utrecht: LOT Dissertation Series 274.

Köhnlein, Björn. 2013. Optimizing the relation between tone and prominence: Evidence from Franconian, Scandinavian, and Serbo-Croatian tone accent systems. *Lingua* 131: 1–28.

Köhnlein, Björn. 2015. An asymmetry in the interaction of pitch and duration. *Diachronica* 32(2): 231–267.

Köhnlein, Björn. 2016. Contrastive foot structure in Franconian tone accent dialects. *Phonology* 33(1): 87–123.

Köhnlein, Björn. 2017. Synchronic alternations between monophthongs and diphthongs in Franconian: A metrical approach. In *Segmental structure and tone*, eds. Wolfgang Kehrein, Björn Köhnlein, Paul Boersma, and Marc Oostendorp, 211–236. Berlin: de Gruyter.

Köhnlein, Björn. 2018. A unified analysis of subtractive morphology in German dialects. *Phonology* 35(4): 617–647.

Köhnlein, Björn. 2019a. Metrically conditioned pitch accent in Uspanteko. In *The study of word stress and accent: Theories, methods and data*. eds. Rob Goedemans, Jeffrey Heinz, and Harry Van der Hulst 293–322. Cambridge: Cambridge University Press.

Köhnlein, Björn. 2019b. Why metrical approaches to tonal accent are worth pursuing. *Phonology* 36(3): 531–535.

Köhnlein, Björn. 2020. Tone accent in North and West Germanic. In *The Cambridge handbook of Germanic linguistics*, eds. Richard Page and Michael Putnam, 143–166. Cambridge: Cambridge University Press.

Köhnlein, Björn, and Yuhong Zhu. 2019. Restricting the power of cophonologies: A representational solution to stem allomorphy in Uspanteko. In *Proceedings of the 2018 annual meeting on phonology*, eds. Katherine Hout, Anna Mai, Adam McCollum, Sharon Rose, and Matt Zaslansky. Washington: Linguistic Society of America.

Krämer, Martin. 2012. *Underlying representations*. Cambridge: Cambridge University Press.

Kuznetsova, Natalja. 2018. Estonian word prosody on the Procrustean bed of morae. *Journal of Estonian and Finno-Ugric Linguistics* 9(1): 209–244.

Lehiste, Ilse. 1965. The function of quantity in Finnish and Estonian. *Language* 41: 447–456.

Lehiste, Ilse. 1997. The structure of trisyllabic words. In *Estonian prosody: Papers from a symposium*, eds. Ilse Lehiste and Jaap Ross, 149–164. Tallinn: Institute of Estonian Language.

McCarthy, John J. 1995. Extensions of faithfulness: Rotuman revisited. Vol. 110 of Rutgers optimality archive. Amherst: Ms., University of Massachusetts.

McCarthy, John J. 2000. The prosody of phase in Rotuman. *Natural Language and Linguistic Theory* 18(1): 147–197.

Mellander, Evan. 2003. (HL)-creating processes in a theory of foot structure. *The Linguistic Review* 20: 243–280.

Morén-Duolljá, Bruce. 2013. The prosody of Swedish underived nouns: No lexical tones required. *Nordlyd* 40(1): 196–248.

Moreton, Elliott. 1999. Non-computable functions in Optimality Theory. Ms., University of Massachusetts, Amherst.

Morrison, Donald Alasdair. 2019. Metrical structure in Scottish gaelic: Tonal accent, glottalisation and overlength. *Phonology* 36(3): 391–432.

Müller, Joseph. 1900. *Untersuchungen zur Lautlehre der Mda. von Aegidienberg*. Bonn: P. Hauptmann'sche Buchdruckerei.

Münch, Ferdinand. 1904. *Grammatik der ripuarisch-fränkischen Mundart*. Bonn: Cohen.

Nörrenberg, Constantin. 1884. Ein niederrheinisches accentgesetz. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 9: 402–412.

Odden, David. 1997. Some theoretical issues in Estonian prosody. In *Estonian prosody: Papers from a symposium*, eds. Ilse Lehiste and Jaap Ross, 165–195. Tallinn: Institute of Estonian Language.

Peters, Jörg. 2006. The Cologne word accent revisited. In *Germanic tone accents. Proceedings of the 1st international workshop on Franconian tone accents*, ed. Michiel de Vaan, *Germanic tone accents. Proceedings of the 1st international workshop on Franconian tone accents*, 107–133. Stuttgart: Franz Steiner Verlag.

Pöchtrager, Markus A. 2006. The structure of length. PhD diss., University of Vienna.

Prillop, Külli. 2013. Feet, syllables, moras and the Estonian quantity system. *Linguistica Uralica* 49(1): 1–29.

Prillop, Külli. 2019. Morae in Estonian. a reply to Natalja Kuznetsova's paper Estonian word prosody on the Procrustean bed of morae. Eesti ja soome-ugri keeleteaduse ajakiri. *Journal of Estonian and Finno-Ugric Linguistics* 10(2): 151–183.

Prince, Alan. 1980. A metrical theory for Estonian quantity. *Linguistic Inquiry* 11: 511–562.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado, Boulder. [Published 2004, Malden, MA and Oxford: Blackwell.]

Ramachers, Stefanie Theodora Maria Rosa. 2018. *Setting the tone. Acquisition and processing of lexical tone in East-Limburgian dialects of Dutch*. Utrecht: LOT.

Rasin, Ezer. 2016. The stress-encapsulation universal and phonological modularity. Talk given at *NAPhC 9*, Concordia University, May 7–8, 2016.

Riad, Tomas, and Jörg Peters. 2021. Northwestern Europe. In *The Oxford handbook of language prosody*, eds. Carlos Gussenhoven and Aoju Chen. Oxford: Oxford University Press.

Schmidt, Jürgen E. 2002. Die sprachhistorische Genese der mittelfränkischen Tonakzente. In *Silbenschnitt und Tonakzente*, eds. Peter Auer, Peter Gilles, and Helmut Spiekermann. 201–233. Tübingen: Niemeyer.

Smith, Jennifer L. 2008. Markedness, faithfulness, positions and contexts: Leniton and fortition in optimality theory. In *Lenition and fortition*. eds. Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 519–560. Berlin: Mouton de Gruyter.

Smith, Laura Catharine. 2020. The role of foot structure in germanic. In *The Cambridge handbook of germanic linguistics*, eds. Michael T. Putnam and B. Richard Page, 49–72. Cambridge: Cambridge University Press.

Spahr, Christopher Edward. 2016. Contrastive representations in non-segmental phonology. PhD diss., University of Toronto.

Topintzi, Nina. 2010. *Onsets: Suprasegmental and prosodic behaviour*, Vol. 125. Cambridge: Cambridge University Press.

van der Hulst, Harry. 2011. Pitch accent systems. In *The Blackwell companion to phonology*, eds. Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice. Vol. 2, 1003–1026. London: Wiley-Blackwell.

van Oostendorp, Marc. 2005. Expressing inflection tonally. *Catalan Journal of Linguistics* 4(1): 107–127.

van Oostendorp, Marc. 2017. Tone, final devoicing and assimilation in Moresnet. In *Segmental structure and tone*, eds. Wolfgang Kehrein, Björn Köhnlein, Paul Boersma, and Marc van Oostendorp, 237–252. Berlin: de Gruyter.

Viitso, Tiit-Rein. 2003. Structure of the Estonian language: Phonology, morphology and word formation. In *Estonian language*, ed. Mati Erelt, 9–92. Tallinn: Estonian Academy Publishers.

Zhu, Yuhong. 2023. A metrical analysis of light-initial tone sandhi in Suzhou Wu. *Natural Language and Linguistic Theory* 41(4): 1629–1678.