The Fabb-Halle approach to metrical stress theory as a window to commonalities between music and language

Laura Dilley[1,2,3] and J. Devin McAuley[2]

Department of Communicative Sciences and Disorders[1]

Department of Psychology[2]

Department of Linguistics[3]

Michigan State University

Address correspondence to:

Dr. Laura C. Dilley

Departments of Psychology and Communication Disorders

Bowling Green State University

Bowling Green, OH 43403

Email: dilley@bgsu.edu

Phone: 419 – 372 – 7182

Fax: 419 – 372 – 6013

To appear in *Language and Music as Cognitive Systems*, P. Rebuschat (ed)., Oxford University Press, 2012.

The world's languages are universally characterized as involving alternations of relatively stronger and weaker prosodic units or events; these alternations of more and less prominent events are often described as rhythmical in nature. The elegant framework for describing these prosodic alternations put forward by Fabb and Halle in Chapter 2 in this volume fits within a general theoretical approach that has come to be known over the last 30 years as *metrical stress theory*. In this chapter, we situate the Fabb-Halle approach within the wider context of metrical stress theory and within linguistics more broadly, while responding to some of the specific claims made in the Fabb and Halle chapter. Taking this broader perspective makes possible an appreciation of the differences between the Fabb-Halle approach to metrical stress theory (hereafter, FH) and other approaches within the same theoretical framework.

The chapter is organized as follows. First, we describe metrical stress theory and how it has evolved over the years. Second, we review claims made by FH and evaluate whether their approach provides a reasonable and better description of basic similarities between music and language than other approaches, such as Lerdahl and Jackendoff (1983), hereafter LJ. Third, we address the descriptive versus explanatory adequacy of the theory, with particular emphasis on the question of whether the theory affords an *explanation* of the various patterns of prosodic prominence and grouping associated with the world's linguistic prosodic systems. Finally, we consider an alternative approach that seeks a perceptual basis for commonalities between music and language.

The origins of metrical stress theory and its role in the study of language

This article aims to provide perspective on (1) how the particular theoretical framework advanced in FH fits into the more general theoretical framework of metrical stress theory, as well as (2) how the specialty area of linguistics represented in FH fits in with the study of language overall. With respect to the former point, FH is one instantiation of metrical stress theory, a theoretical approach which has sought generally to account for word- and phrase-level stress patterns and other prosodic phenomena in terms of hierarchical constructs involving metrical grid and tree structures similar to those used in music to account for relative prominence. The seminal ideas behind metrical stress theory were first advanced by Liberman (1975) in his Ph.D. dissertation; since then many variations on the central ideas of this original work have been put forward. Liberman's proposals, elaborated in Liberman and Prince (1977), centered on the idea that variations in stress level are due to underlying hierarchical structures of prominence or grouping. Liberman's proposals provided a counterpoint to the dominant theoretical approach at the time, which was to treat stress as a type of segmental distinctive feature. (See Jakobson, Fant, & Halle, 1952; Stevens, 2000 for a discussion of distinctive features.) The treatment of stress as a distinctive feature had built on the work of the American structuralist school of linguistics (e.g., Trager & Smith, 1951) and had culminated in *The Sound Pattern of English* (Chomsky & Halle, 1968), often abbreviated *SPE*. Liberman's insight was that variations in word stress can be explained by hierarchical structures (e.g., grid structures) similar to those found in music. The observations afforded under Liberman's approach made word stress appear much more regular than it seemed under competing accounts and obviated the need for assuming linguistic mechanisms such as cyclic rule application and stress subordination proposed in *SPE*.

The inherent connection between music and language has remained a running theme throughout much, but not all, subsequent work that has modified and refined the original ideas behind metrical stress theory. Significant debates in the 1980's centered on the relative importance for linguistic accounts of prosodic phenomena of two sorts of interrelated mechanisms for accounting for metrical prominence introduced in Liberman (1975). The first was the metrical grid, in which discrete events (e.g., syllables) were represented by X's; these X's indicate time slots, and the degree of prominence is represented by the height of a column of X's. The second was a tree structure, an example of which is shown in Figure 1. The nodes of a tree are labeled *s* (strong) and *w* (weak) to mark relative prominence, and grids (patterns of relative prominence) can then be read off the tree.

Early work in metrical theory (e.g., Hayes, 1984; Liberman, 1975; Liberman & Prince, 1977) emphasized tree formalisms while also making use of grids; many subsequent arguments centered on whether both grids and trees were necessary to account for prominence patterns and related phenomena across languages. Noteworthy proponents of a 'grid-only' approach were Prince (1983) and Selkirk (1984), while the charge for a 'tree-only' approach was led by Giegerich (1985). Hybrid approaches of various sorts incorporated both a mechanism for grouping (e.g., trees), as well as a mechanism for relative prominence (i.e., a grid of some kind) (Halle & Vergnaud, 1987; Hammond, 1984; Hayes, 1983, 1984) . Hammond (1984) and Halle and Vergnaud (1987) in particular developed a structure that has come to be called a 'bracketed grid', in which X's are grouped by parentheses; this type of structure survives to populate the FH paper.

Connections between music and language have played a role in various instantiations of metrical theory to varying degrees, from not at all to substantial.

The second consideration which provides perspective on FH is that this work and almost all of the work on metrical stress theory falls within a specialty branch of linguistics known as *theoretical phonology*. Researchers working in this area aim to develop broad, theoretical accounts for similarities and differences in the sound patterns of the world's languages, focusing largely on theoretical constructs and linguistic abstractions. Many theoretical phonologists do not focus on details of sound signals or properties of human perception and cognition which might shape linguistic knowledge and intuition; experiments involving manipulation of sound stimuli also fall largely outside of theoretical phonology.[1] Given that many sub-disciplines of linguistics and psychology deal in some way with the study of language-based sound systems, it is noteworthy that certain statements made in FH apply narrowly to the sub-discipline of theoretical phonology, for example: "It is the goal of the linguist to discover which combination of grid-building rules, and which stress-assigning rules, are required to explain the pattern of stresses which we find in the words of a specific language." In this statement, FH are clearly equating "linguist" with "theoretical phonologist"; researchers working in other sub-disciplines might well state the goals of their inquiries into sound systems differently.

Having considered the history of metrical stress theory and the field of linguistics generally, it is now possible to better appreciate particulars of the FH approach. The account proposed in FH permits a description of variations in stress systems of many languages using a combination of bracketed grids, iterative rules, and other types of rules.

FH claim that this theoretical approach is simpler than that of LJ, who propose a series of metrical and grouping well-formedness rules which act to rule out impossible grid structures. Moreover, FH claim that their theory makes music and language appear closer than the work of LJ. FH provides an elegant description of stress patterns, and its ability to account for data from a variety of languages, as well as different types of structure (e.g., prose vs. poetry), is provocative. In the remainder of the chapter, we consider two issues. First, to what extent does FH provide a reasonable description of basic similarities between music and language? Second, to what extent does the theory afford an explanation of the various patterns of prosodic prominence and grouping associated with the world's linguistic prosodic systems?

The FH approach as a reflection of the "closeness" of music and language

Does the FH account entail a description that reflects the basic similarities between music and language? Stated differently, does FH's claim that their theory makes music and language appear closer than LJ's stand up? In the following we consider three ways in which the FH instantiation of metrical stress theory presents a view of language that makes it seem relatively more dissimilar to music than might be otherwise envisioned.

First, consider that in FH and in many other versions of metrical theory, strong and largely *a priori* restrictions are assumed on types of grouping structures that can be formed from syllables in a word, leading to the appearance of greater dissimilarity between linguistic and musical structures than if these restrictions were not assumed. For example, FH assume that parses of metrical events on a given line of the grid must be all-binary or all-ternary. That is, for a given line of the grid, parsing rules apply iteratively such that all (complete) groups have exactly the same number of syllables (two or three). In other words, a group of two can't occur adjacent to a group of three or vice versa. This presents a view of language which suggests deviation from music in several respects. In music, there is no prohibition against alternations of binary and ternary groupings – groupings of two can follow groupings of three without penalty.

Perhaps of more concern for whether metrical theory reflects similarities in music and in language is the fact that the assumption in FH of all-binary or all-ternary parses entails an adjunct assumption of something called *extrametricality*. Extrametricality is assumed to cause a syllable not to belong to *any* group; such a mechanism is useful in theories which seek to limit the number of possible grouping structures by permitting rules for stress assignment to "skip over" certain syllables in counting, treating them as if they were invisible. The notion of extrametricality has a long and rich history in theoretical phonology (Halle & Vergnaud, 1987; Hayes, 1979, 1981; Ito, 1986; Liberman & Prince, 1977; Roca, 1992); see Hayes (1995, pp. 58-60) for an overview. To illustrate how this mechanism works, consider that the word *extrametricality* seems to surface as a sequence of two binary feet plus a ternary foot, i.e., (s w) (s w) (s w w); this word would be parsed under FH and related approaches as three binary feet plus a final extrametrical syllable, i.e., (s w) (s w) (s w) w. A notion of extrametricality is critical to the FH account and much related work in metrical stress theory. While FH assumes that ternary feet exist in the inventory of possible structures, other so-called parametric versions of metrical theory take the hard line of assuming that ternary feet are not part of the possible inventory of grouping structures across languages. For example, Hayes (1981; 1995) dispenses with ternary feet entirely, instead assuming the existence of only binary (or unbounded) feet. Under associated supporting assumptions, Hayes accounts for attested ternary linguistic stress systems (i.e., languages in which stresses occur every three syllables within words, including Cayuvava, Pacific Yupik, and Sentani), as repeated sequences of a binary foot, plus an extrametrical syllable. (See Hayes, 1995, Ch. 8, for a discussion of these systems.) While these parametric versions of metrical theory comprise elegant descriptions with clever mechanisms for deriving stress patterns of words in many languages, the constraining nature of parametric assumptions of all-binary or all-ternary parses, combined with notions of extrametricality, seem to imply a wider divide between language and music than in an alternative theory which does not place such restrictions on ternary groupings.

In music, there is no such thing as extrametricality, and elements in a closely connected musical sequence do not occur "ungrouped". Rather, closely connected musical elements tend to be heard as grouped into twos and threes. (See Handel, 1989 for a review.) If fewer restrictions were placed on foot inventories of languages in under parametric metrical theories (e.g., the assumption that parses are all-binary or all-ternary), there would be no clear need for an adjunct notion of extrametricality, at least to account for prosodic phenomena.[2] These restrictions do not occur in music, and LJ do not assume *a priori* all-ternary or all-binary parses, nor a notion of extrametricality, to account for musical data; hence, the LJ approach seems to present a window in which similarities between music and language appear closer than under FH and related linguistic work.

A second respect in which the FH approach seems to cast a wider chasm between music and language has to do with the mechanism(s) associated with parenthesis insertion. The proposed mechanism in language for meter and grouping is an "iterative" rule which inserts either a left or right parenthesis at regular intervals (every 2 or every 3 syllables). Such an iterative grouping mechanism (consisting of insertion of either right or left parentheses in sequences of abstract timing events) does not have any clear parallel in music. In essence, the mechanism assumes two different types of juncture ("right" and "left"). The distinction in parenthesis type, together with the ability to count iteratively from either the left or right side of the word, permits the theory to render extrametrical a syllable either at the beginning or at the end of a string. But what does a distinction in type of parenthesis mean? What is the difference between a juncture of a "left" parenthesis vs. a "right" parenthesis? It is not clear what these correspond to in terms of phonetic characteristics, perceptual or structural conceptualizations, etc. The explanatory power of these distinctions is unclear when considered in a broader context. Again, the distinction in parenthesis type, together with the notion of extrametricality, are unnecessary if it is simply assumed that groups can be a mixture of binary and ternary feet.

Third, there are problems with the notion of "iterativity". The iterative application of grouping of events into twos and threes is a cornerstone of the FH account and related work (e.g., Idsardi, 1992). Yet, there is no clear corresponding notion of iterativity in music. Iterativity might or might not be likened to "hysteresis" – however, the strictness of structure implied by the iterativity mechanism is not the case for music. Variations in timing, accentuation, etc. are readily adapted to, undercutting a strict notion of "iterativity" in music. FH note that the notion of iterativity is controversial from the standpoint of linguistics as well, in that iterative rules do not otherwise occur in linguistic or musical sound systems . (See also Halle, 1990; Halle & Vergnaud, 1987.)

Descriptive versus explanatory adequacy of the FH approach

It is important to consider what the FH approach does and does not attempt to do. In this regard, while the FH approach describes the data, it does not seek to explain how linguistic communication by sound works. Indeed, it considers only a limited set of structures which humans arguably might be thought to have in their heads. Moreover, it does not seek to describe, nor to explain, how differences of sound structure (e.g., the loudness level of syllables) might physically convey prosodic differences, nor how listeners might interpret these differences of detail of the speech signal in terms of underlying structures. FH assume that their system operates at a wholly separate level of description from the principles which are operational for physical, linguistic sound systems. Moreover, it seems unreasonable that FH should be held to the standard of explaining nuances of perception, and how sound systems convey prosodic grouping and prominence through details of physics. However, there are compelling reasons to consider whether other types of explanations might afford greater, or at least complementary, insight into how speakers and listeners communicate the nuances of prosodic properties in language.

One noteworthy aspect of the FH view and much other work in metrical phonology is the explicit assumption that linguistic structures making up grids are not privy to introspection. Moreover, FH assume that individuals' perceived sense of grouping may or may not correspond to underlying linguistic groups. By taking this tack FH assume that behavioral judgments and other types of experimentation aimed at uncovering the structures which individuals have in their heads bear no definable relationship to the types of structures in their metrical theory. As a result, the FH approach is moved into a realm in which it cannot reasonably be tested by most, if not all, behavioral methods. The assumption of a lack of connection between listeners' surface perceptions and behavior and the underlying linguistic structures leads to an increase in complexity in any overarching theory of linguistic communication via prosodic sound systems. This is because such an approach requires a wholly separate mechanism to be posited to explain *perceived* groupings, as separate from the underlying ones. A theory which proposes (or implicitly assumes) two types of grouping – one implicit and the other explicit, where the implicit (linguistic) grouping might have an inherently different structure that the explicit grouping, is not very parsimonious.

A common perceptual basis for musical structure and prosodic patterns in language

An alternative approach which is being pursued by a number of researchers is to assume as a starting point the likelihood of shared mechanisms for processing music and language which might be responsible for both perception and production of metrical patterns in both domains. For example, work by Dilley (2005; 2008) builds on traditional metrical stress frameworks and facts about English intonation while positing a strong connection between music and language. Other recent work (Dilley & McAuley, 2008) aims to identify whether specific properties of sound (amplitude, frequency, duration, timbre, etc.) lead to similar perceptions about grouping and meter in speech and non-speech auditory perception, a finding which would support shared processing mechanisms for speech and non-speech perception. Our work builds on a relatively large literature on non-speech auditory perception illustrating effects of frequency, duration, and amplitude patterning on perceived organization of auditory sequences. In general, when individuals hear simple tone sequences, the frequency, duration and amplitude patterning of sequence elements (i.e., tones) conveys a sense of sequence organization and structure. Perceived organization includes the sense that some sequence elements belong together (i.e., they are grouped), that within a group some elements are accented, while others are not, and that accent patterns tend to repeat. For example, in an isochronous sequence of tones of equal amplitude and duration alternating between a fixed high (H) and fixed low (L) frequency, e.g., HLHLHL, listeners tend to hear repeating strong-weak binary groupings of tones with either the high or low tone as accented and beginning each group, i.e., (<u>H</u>L)(<u>H</u>L)(<u>H</u>L) or (<u>L</u>H)(<u>L</u>H)(<u>L</u>H) (Woodrow, 1909, 1911). Similarly, repeating strong-weak binary patterns of accents induced by distal (i.e., remote, distant or nonlocal) frequency, duration, and/or amplitude patterning of sequence elements tends to generate periodic expectations about the grouping and perceived accentuation of later sequence elements, even when there are no explicit proximal (i.e., local) acoustic cues to grouping and accents in those elements (Boltz, 1993; Jones, 1976; Jones & Boltz, 1989; Large & Jones, 1999; McAuley & Jones, 2003; Povel & Essens, 1985; Thomassen, 1982).

Our experiments used lexically ambiguous syllable sequences (e.g., *footnote bookworm, foot notebook worm*) to examine how distal frequency and timing cues would affect segmentation of these proximal syllable strings into words. The acoustic characteristics of the final three syllables in our target experimental strings were held constant with H, L and H pitch, and the pitch and durational pattern of the preceding, distal syllables was manipulated using computer speech resynthesis techniques. We showed that perception of grouping of syllables into words depended on the preceding prosodic context in precisely the way expected if listeners were applying principles to speech that they apply to nonspeech auditory sequences alternating in pitch and/or duration (Dilley & McAuley, 2008), even though the acoustic characteristics of the syllables were identical. Support was found for a perceptual grouping hypothesis in which distal prosodic characteristics established perceived patterns of pitch and rhythm that affected the grouping of syllables into prosodic units in just the manner expected for nonspeech auditory perception, thereby influencing word segmentation and lexical processing. This work fits into a growing body of interdisciplinary findings demonstrating evidence for shared processing resources in music and in language. We feel that such interdisciplinary experimental lines of work are likely to be profitable in terms of coming to understand how and why speakers perceive and produce prosodic patterning as they do in language. In this regard, it seems useful to pursue interdisciplinary research agendas in which experimental techniques as well as descriptive linguistic apparatuses such as those afforded in FH and related work can be brought to bear in understanding linguistic processing.

A variety of approaches are being pursued to uncover evidence of common processing mechanisms for music and speech. For example, Patel and colleagues are actively pursuing the question of shared processing mechanisms for music and language using a wide range of techniques. These include investigations of pitch, timing, and structural processing deficits in individuals with acquired or congenital amusia (Patel, 2005; Patel, Foxton, & Griffiths, 2005; Patel, Iversen, Wassenaar, & Hagoort, 2008), comparisons of human and nonhuman rhythm perception and production ability (Patel & Iversen, 2006; Patel, Iversen, Bregman, Schulz, & Schulz, 2008), investigations of processing demands in normal speakers using standard psycholinguistic tasks (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Slevc, Rosenberg, & Patel, 2008), and comparisons of temporal and pitch properties of speech and music from different cultures (Patel & Daniele, 2003; Patel, Iversen, & Rosenberg, 2006). Moreover, additional work aims to identify production constraints on rhythmic sequences in speech which affect both linguistic and musical performance similarly, while building alternative linguistic theoretical frameworks to accommodate such findings (Cummins & Port, 1998; Port, 2003). Perspectives on music and language processing are also greatly aided by developmental work (Hannon & Johnson, 2005; Hannon & Trehub, 2005) and by use of electrophysiological measures comparing processing of speech and music (Magne, Schon, & Besson, 2003; Snyder & Large, 2005). We can envision the outline of a theory which builds on many core insights of metrical stress theory, including those embraced in FH, but in which principles of perception and cognition would play a central role to jointly explain facts about music and language. In this respect, we agree with FH that research is presently far from this goal, but that it is a desirable one to try to achieve. To this end, a theoretical perspective which both strives for parsimony and holds high the aim of reflecting commonalities between music and language seems most likely to afford the greatest insight into sound-based human communication systems.

References

Boltz, M. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception and Psychophysics, 53*(6), 585-600.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Cummins, F., & Port, R. F. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*, 145-171.

Dilley, L. C. (2005). *The phonetics and phonology of tonal systems.* Unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Dilley, L. C. (2008). On the dual relativity of tone. In *Proceedings of the 41st Meeting of the Chicago Linguistics Society (2005)* (Vol. 41, pp. 129-144). Chicago, IL.

Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language, 59*(3), 294-311.

Fedorenko, E., Patel, A. D., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory and Cognition, 37*(1), 1-9.

Giegerich, H. J. (1985). *Metrical Phonology and Phonological Structure: German and English*. Cambridge: Cambridge University Press.

Halle, M. (1990). Respecting metrical stucture. *Natural Language and Linguistic Theory, 8*, 149-176.

Halle, M., & Vergnaud, J.-R. (1987). *An Essay on Stress*. Cambridge, MA: MIT Press.

Hammond, M. (1984). *Constraining metrical theory: A modular theory of rhythm and destressing.* Unpublished Ph.D., University of California, Los Angeles.

Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events*. Cambridge, MA: MIT Press.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology, 50*, 354-377.

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences (USA), 102*, 12639-12643.

Hayes, B. (1979). Extrametricality. *MIT Working Papers in Linguistics, 1*, 77-86.

Hayes, B. (1981). *A metrical theory of stress rules.* Unpublished Ph.D., Massachusetts Institute of Technology, Cambridge, MA.

Hayes, B. (1983). A grid-based theory of English meter. *Linguistic Inquiry, 14*, 357-394.

Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry, 15*, 33-74.

Hayes, B. (1995). *Metrical Stress Theory*. Chicago: University of Chicago Press.

Idsardi, W. (1992). *The computation of prosody.* Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Ito, J. (1986). *Syllable theory in prosodic phonology.* Unpublished Ph.D., University of Massachusetts, Amherst.

Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to Speech Analysis*. Cambridge, MA: MIT Press.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323-355.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459-491.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119-159.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Liberman, M. (1975). *The intonation system of English.* Unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*(2), 249-336.

Magne, C., Schon, D., & Besson, M. (2003). Prosodic and melodic processing in adults and children: Behavioural and electrophysiological approaches. *Annals New York Academy of Sciences, 999*, 461-476.

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception & Performance, 29*(6), 1102-1125.

Patel, A. D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Annals of the New York Academy of Sciences, 1060*, 59-70.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87,* B35-B45.

Patel, A. D., Foxton, J., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminiating intonation contours extracted from speech. *Brain and Cognition, 59*, 310-313.

Patel, A. D., & Iversen, J. R. (2006). A non-human animal cam drum a steady beat on a musical instrument. In M. Baroni, A. R. Adessi, R. Caterina & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception & Cognition* (pp. 477). Bologna, Italy.

Patel, A. D., Iversen, J. R., Bregman, M. R., Schulz, I., & Schulz, C. (2008). *Investigating the human-specificity of synchronization to music.* Paper presented at the 10th International Conference on Music Perception and Cognition, Sapporo, Japan.

Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *Journal of Acoustical Society of America, 119*, 3034-3047.

Patel, A. D., Iversen, J. R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology, 22*, 776-789.

Pierrehumbert, J., Beckman, M., & Ladd, D. R. (2000). Conceptual foundations of phonology as a laboratory science. In N. Burton-Roberts, P. Carr & G. J. Docherty (Eds.), *Conceptual and Empirical Foundations of Phonology* (pp. 273-303): Oxford University Press.

Port, R. F. (2003). Meter and speech. *Journal of Phonetics, 31*, 599-611.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*(4), 411-440.

Prince, A. (1983). Relating to the grid. *Linguistic Inquiry*, *14*, 19-100.

Roca, I. (1992). Constraining Extrametricality. In *Phonologica 1988*. Cambridge: Cambridge University Press.

Selkirk, E. O. (1984). *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge, MA: MIT Press.

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2008). Language, music and modularity: Evidence for shared processing of linguistic and musical syntax. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima & M. Tsuzaki (Eds.), *10th International Conference on Music Perception and Cognition* (pp. 598-605). Sapporo, Japan.

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*, 117-126.

Stevens, K. N. (2000). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Thomassen, J. M. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America*, *71*(6), 1596-1605.

Trager, G. L., & Smith, H. L. (1951). *An outline of English structure*. Washington, DC.

Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology, 14*, 1-66.

Woodrow, H. (1911). The role of pitch in rhythm. *The Psychological Review, 18*, 54-77.

Footnotes

[1]Note that the distinct but related sub-area of linguistics known as *laboratory phonology* integrates theoretical work with empirical data from experiments. See Pierrehumbert et al. (2000) for more information on this approach.

[2] The notion of extrametricality has also been applied to segments (i.e., phonemes) comprising a syllable, in addition to syllables themselves. For example, extrametricality is invoked to explain certain facts about syllabification, i.e., how segments are arranged into syllables (Ito, 1986)..Such arguments are often brought up in support of the notion of extrametricality generally, thereby bolstering its role in prosodics.

Figure captions.

Figure 1. An example of a tree structure used in some versions of metrical stress theory.

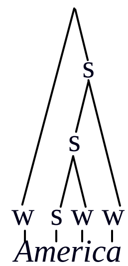

Figure 1.