Morpheme-internal Recursion in Phonology

Studies in Generative Grammar

Volume 140

Morpheme-internal Recursion in Phonology

Edited by
Kuniya Nasukawa

DE GRUYTER
MOUTON

ISBN 978-1-5015-1777-8
e-ISBN (PDF) 978-1-5015-1258-2
e-ISBN (EPUB) 978-1-5015-1241-4
ISSN 0167-4331

Library of Congress Control Number: 2019957104

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: Integra Software Services Pvt. Ltd.
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Contents

List of Contributors

Phillip Backley is Professor of English Linguistics at Tohoku Gakuin University, Japan. He gained a PhD in phonology from the University of London (UCL) in 1998, and has published on a range of topics in theoretical phonology, language acquisition and the history of English. He is actively involved in the development of an element-based (cf. feature-based) approach to segmental representation, and is the author of *An Introduction to Element Theory* (EUP, 2011) and co-editor (with Kuniya Nasukawa) of *Strength Relations in Phonology* (Mouton 2009).

Edoardo Cavirani is a FWO Marie Skłodowska-Curie Actions—Seal of Excellence postdoctoral researcher at the Center for Research in Syntax, Semantics and Phonology of KU Leuven. He holds a MA in Historical and Comparative Linguistics from the University of Pisa and a PhD (Doctor Europaeus) from the University of Pisa and the University of Leiden.

Marcel den Dikken is Senior Researcher at the Theoretical Linguistics Department of the Research Institute for Linguistics) of the Hungarian Academy of Sciences, Budapest, and Research Professor at the School of English and American Studies of Eötvös Loránd University (ELTE). He is series editor (with Liliane Haegeman, Joan Maling and Maria Polinsky) of *Studies in Natural Language and Linguistic Theory* (Springer, Dordrecht), editor of *The Cambridge Handbook of Generative Syntax* (CUP 2013), author of *Particles: On the Syntax of Verb-particle, Triadic, and Causative Constructions* (OUP 1995), *The Structure of the Noun Phrase in Rotuman* (LINCOM Europa 2003), *Relators and Linkers: The Syntax of Predication, Predicate Inversion, and Copulas* (MIT Press 2006) and *Dependency and Directionality* (CUP 2018), and co-author (with Hans Broekhuis) of *Syntax of Dutch. Nouns and Noun Phrases, Vol. 2* (Amsterdam University Press 2012).

Harry van der Hulst is Professor, Linguistics at University of Connecticut. His research interests cover various aspects of phonology, with a focus on word and phrasal stress, syllabic structure, features and segmental structure, sign language, language change and evolution, the interplay between phonetics and phonology, and linguistic analysis of graphic narratives. He has published over 170 articles and 29 edited volumes in these areas. Recent work includes *Asymmetries in Vowel Harmony: A Representational Approach* (2018, Oxford University Press) and *Radical CV Phonology: A Theory of Segmental and Syllabic Structure* (2020, Edinburgh University Press). Van der Hulst is Editor-in-Chief of *The Linguistic Review* and co-editor of the series 'Studies in generative grammar' (both with De Gruyter Mouton).

Nancy C. Kula is Professor of linguistics in the Department of Language and Linguistics at the University of Essex. She has worked on a range of issues in phonology including segmental representations, tone and prosody and phonological interfaces with morphology and syntax. Her empirical focus mainly centers on Bantu languages. She is co-editor (with Bert Botma and Kuniya Nasukawa) of the *Bloomsbury Companion to Phonology* and has published in a wide range of journals.

Chihkai Lin received his M.A. degree from the Graduate Institute of Linguistics at National Taiwan University (2008) and his Ph.D. degree from the Department of East Asian Languages and Literatures at the University of Hawaiʻi at Mānoa (2015), where he studied under

https://doi.org/10.1515/9781501512582-203

Dr. Alexander Vovin. Currently, he is a project assistant professor in the Department of Applied Foreign Languages at Tatung University (Taiwan). He specializes in phonology and historical linguistics focusing on East Asian languages. His research interests include historical changes in Chinese dialects, Japanese and Korean based on philological data, and synchronic phonology in Austronesian languages, Chinese, Japanese and Korean.

Xiaoxi Liu is a final year PhD student in the Department of Language and Linguistics at the University of Essex, working on Element Theory representations with focus on depressor effects and broader consonant-tone interaction. Her thesis investigates the similarities and differences between depressors in Bantu languages and Chinese dialects with the goal of proposing a unified analysis that exploits multi-layered element representations.

Filiz Mutlu is a graduate student of linguistics at the University of Ljubljana. She started developing a theory of phonological representations in 2015 at Boğaziçi University for her MA thesis, with the central idea that asymmetry is the building principle of phonology as well as other components of language. Currently she is working on her doctoral thesis, where she develops her theory (Spatial Phonology) further and claims that linguistic structures are geometric objects of up to three dimensions. These objects are roughly translatable to trees, but crucially they have underpinnings in the representation of space and magnitude in vertebrates structures grow to the right, upwards and to the front.

Kuniya Nasukawa is Professor of English Linguistics at Tohoku Gakuin University, Japan. He has a Ph.D. in Linguistics from University College London (UCL), and his research interests include prosody-melody interaction and precedence-free phonology. He has written many articles covering a wide range of topics in phonological theory. He is author of *A Unified Approach to Nasality and Voicing* (Mouton 2005), co-editor (with Phillip Backley) of *Strength Relations in Phonology* (Mouton 2009), co-editor (with Nancy C. Kula and Bert Botma) of *The Bloomsbury Companion to Phonology* (Bloomsbury 2013), and co-editor (with Henk van Riemsdijk) of *Identity Relations in Grammar* (Mouton 2014).

Hitomi Onuma is a full-time lecturer at Iwate Medical University, Japan. She received her Ph.D. in Literature from Tohoku Gakuin University in 2015. Her research interests and publications include various aspects of segmental and prosodic phonology, particularly concerning the status and behaviour of empty nuclei. Currently she has worked in the framework of precedence-free phonology.

Marc van Oostendorp is Senior Researcher at the Department of Variationist Linguistics at the Meertens Institute of the Royal Netherlands Academy of Arts and Sciences. He holds an MA in Computational Linguistics and a PhD from Tilburg University. He is co-editor (with Colin J. Ewen, Elizabeth V. Hume and Keren Rice) of *The Blackwell Companion to Phonology* (Wiley-Blackwell 2011).

Markus Pöchtrager received his Ph.D. in General Linguistics at the University of Vienna in 2006. From 2008 to 2017 he was a faculty member at the department of linguistics at Boğaziçi University (İstanbul), and from 2017 onwards at the University of Vienna. His research focuses on phonology, phonological metatheory, phonology-morphology interaction, linguistics as a cognitive science, Finno-Ugric and Scandinavian languages as well as Altaic and Caucasian languages.

Clemens Poppe is an Assistant Professor of Linguistics at Waseda University, Japan. He holds a Ph.D. in Linguistics from the University of Tokyo, and previously worked as a postdoctoral researcher at the National Institute for Japanese and Linguistics in Tokyo under the auspices of the Japan Society for the Promotion of Science. His fields of interest include phonology, morphology, and dialectal variation.

Geoff Schwartz is Associate Professor in the Faculty of English at Adam Mickiewicz University (UAM), Poznań, Poland. He received his Ph.D. in Slavic Linguistics from the University of Washington (Seattle) in 2000. He has been employed at UAM since 2002, where his teaching has concentrated on acoustic phonetics and L2 speech acquisition, and was awarded a post-doctoral degree (Polish habilitacja) in 2010. In recent years, he has been working in phonological theory, developing the Onset Prominence representational framework, and continuing his work on L2 phonological acquisition. His articles have appeared in numerous journals, including *Journal of Linguistics*, and he is a regular participant at important phonology and L2 acquisition conferences.

Ali Tifrit is Associate Professor of Linguistics at the University of Nantes, France. He is member of the Laboratoire de Linguistique (LLing UMR6310 CNRS—University of Nantes). His research interests are focusing on formalisms and on the definition of phonological primitives (features and elements) and their relation to phonological activity. He has a particular interest in the synchrony and diachrony of French and Romance languages. He published works on the hierarchical organisation of segments in a unary perspective and on phonological processes (palatalization, spirantization, epenthesis) in a Government Phonology framework.

Hisao Tokizaki is Professor of English Linguistics at Sapporo University, Japan. He has a Ph.D. in Linguistics from University of Tsukuba, and his research interests include the syntax-phonology interface and linguistic typology. He has written many articles covering a wide range of topics in the theory of interface. He is the author of *Syntactic Structure and Silence: A Minimalist Theory of Syntax-Phonology Interface* (Hitsuji syobo), co-editor (with Yoshihito Dobashi) of a special issue of *Linguistic Analysis* 38(3/4), *Universal Syntax and Parametric Phonology* (2013).

Kuniya Nasukawa

Introduction

Motivation

For the Generative Grammar approach to the study of language, recursion is a key concept not only in the debate on how structure is organised within the language faculty but also on the origins of language itself. Researchers with an interest in the evolution of language (Reuland 2009, Chomsky 2010) have argued that natural language in its present form developed from single-word expressions through the use of a recursive merge device.

The phonology literature also often refers to recursive structure in the analysis of recurrent phenomena such as stress and intonation patterns, which rely on (morpho-)syntactic structures generated by syntactic computation. As such, this recursive structure is in fact (morpho-)syntactic rather than phonological (Scheer 2011, Nasukawa 2015).

To establish whether recursive structure exists in phonology or not, we must investigate recursion in morpheme-internal phonological structure, since this cannot be accessed by (morpho-)syntax. There are two opposing views about recursion in phonology.

One view maintains that phonology is recursion-free (no recursive structure-building). Instead, it assumes that phonological structure within a morpheme consists of a set of linearly ordered segments (a string-based flat structure) in the lexicon, and that phonology is not responsible for building lexical phonological structure (Pinker and Jackendoff 2005; Neeleman and van de Koot 2006; Samuels 2009; Scheer 2008, 2011).

According to the opposing view, morpheme-internal phonological structure refers not to precedence properties but to a set of features which are concatenated hierarchically. In this approach, syntax is responsible for merging not only lexical items such as morphemes and words, but also phonological categories (primitives) to build the phonological structure of morphemes in the lexicon (Nasukawa 2014, 2015, 2016; Nasukawa and Backley 2015).

This volume will provide the first platform for debate on the place of recursive structure in phonology and on the formal status of phonology in the language faculty. It has its origins in the workshop entitled 'Recursion in Phonology' held at Tohoku Gakuin University, Sendai, Japan on 1–2 September 2016, where six of the papers included here were first presented (Chihkai Lin, Kuniya Nasukawa, Hitomi

https://doi.org/10.1515/9781501512582-001

Onuma, Clemens Poppe, Geoffrey Schwartz and Hisao Tokizaki). The remaining papers are by other prominent scholars in the field. It is encouraging that the issue of recursion in phonology has also attracted interest and support from those leading the research project 'Evolinguistics: Integrative Studies of Language Evolution for Co-creative Communication' (funded by MEXT/JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas #4903, Grant Number JP18H05081).

Clearly, a volume of this size cannot do justice to a topic as broad as that of recursion in morpheme-internal phonology. Nevertheless, we hope that these papers will convey something of the scope and influence that recursive hierarchical structure appears to have on the analysis of apparently unrelated phenomena across different languages and different domains of linguistic study.

Abstracts

Phillip Backley and Kuniya Nasukawa
Recursion in melodic-prosodic structure

Backley and Nasukawa argue that Phonological information comes in two kinds: melodic information describes the qualitative properties of sounds while prosodic information describes how sounds are organised into larger structures such as syllables and feet. Traditionally, each is represented as an independent domain and described using a unique set of structural units. But there are advantages in integrating the two domains into one unified structure. In this paper, melody-prosody integration succeeds by allowing elements, the units of melody, to also function as prosodic constituents, thereby eliminating the need for labels such as *nucleus*, *syllable* and *foot*. The smallest prosodic domain ('nucleus') is represented by an element from the set {|A|, |I|, |U|}, chosen by parameter to reflect the quality of the default vowel in a language, e.g. English has |A| as its structural head to reflect its default [ə]. Contrastive vowels are then expressed by allowing the head to support dependent structure – constructed via the recursive concatenation of elements functioning as units of melody. Reversing established assumptions, dependents are the main contributors of linguistic (contrastive) information while heads take on a largely structural role. This brings phonology more into line with syntax, where dependents rather than heads are informationally rich.

Edoardo Cavirani and Marc van Oostendorp
A theory of the theory of vowels

Cavirani and van Oostendorp represent the most common vowel contrasts in a theory that allows only (recursive) embedding of treelets. Such a theory needs neither features nor elements. We show that from such a theory we can actually derive some common properties of the element set |A, I, U|: why are there only three of them? And why does |A| behave differently from the other two? Furthermore, the theory also gives a natural place to both schwa and the completely empty nucleus. We also show how this theory is related to some earlier proposals in the literature.

Marcel den Dikken and Harry van der Hulst
On some deep structural analogies between syntax and phonology

The principal aim of **den Dikken and van der Hulst's** chapter is to bring phonology and syntax together with an outlook on linguistic analysis that uses the same representational system in morphosyntax and all levels of phonological analysis, including phonological structure above the syllable, the internal organization of the syllable, and the structure of segments. The central tenet of the approach is the generalization of complementation, specification, adjunction and conjunction relations from syntax to phonological structure. Recursive X-bar-theoretic structures are employed in phonology in the representation of geometrical relations of all kinds (both segmental and suprasegmental). A special role in the phonosyntax of the syllable/foot is played by the phonological counterpart to the 'light *v*' of syntactic structures. The chapter closes by offering an explanation for the fact that recursion in phonology is less pervasive than it is in syntax.

Chihkai Lin
Decomposition and recursive structure: Glide formation and vowel lowering in East Asian languages

Lin argues that a sequence of Sino-Japanese vowels [e] and [u] undergoes glide formation and vowel lowering ([eu] → [jo]). A similar sequence is attested in Tsou, but only glide formation occurs ([eu] → [ju] or [eu] → [ew]).

In Sino-Korean, a sequence of vowels [o] and [i] also undergoes glide formation and vowel lowering ([oi] → [we]). The objective of this study is to investigate glide formation and vowel lowering in the three languages from an element-based approach, paying specific attention to the necessity of decomposition and the application of recursive structure. It is shown that Sino-Japanese mid vowel [e] is decomposed into two elements |I A| and Sino-Korean mid vowel [o] into elements |U A|. The decomposed element |I| or |U| undergoes glide formation. In addition, if the decomposed element |A| interacts with the following high vowel, the high vowel is lowered. In Tsou, vowel lowering is not attested. To differentiate the changes in Sino-Japanese and Sino-Korean from the processes in Tsou, Lin suggests that recursive structure is an inevitable mechanism for the changes, [eu] → [jo] in Sino-Japanese and [oi] → [we] in Sino-Korean. In Tsou, recursive structure is necessary for the process, [eu] → [ew].

Xiaoxi Liu and Nancy C. Kula
Multi-layered recursive representations for depressors

Liu and Kula investigate depressor consonant effects as an example of the interaction between segmental and prosodic structure, in particular consonantal structure and tone. Contrary to expectation it is shown that the whole spectrum of laryngeal specifications can trigger depressor effects viz. voicing, breathiness, (voiceless) aspiration and plain voiceless, within the range of predominantly southern Bantu languages investigated. This distribution is accounted for by proposing a multi-layered recursive element geometry that allows the element |L| – central to the representation to depressor effects – to be represented recursively on different levels in a hierarchical representation with the flexibility of |L| appearing in different dominance relations that then allow the different laryngeal specifications to act as depressors. The connection between element |L| and depression follows from the tripartite identity of |L| in Element Theory as representing voicing, low tone, and nasality. The proposed recursive structure captures the complex depressor effects and at the same time manages to account for the asymmetry between, on the one hand, attested low tone – voicing interaction and, on the other hand, unattested low tone – nasality interactions.

Filiz Mutlu
Embedding of the same type in phonology

Mutlu models consonant clusters and affricates as recursive structures within a novel theory. The view of phonological structure offered here is basically identical to syntactic structure. That is, consonants are represented as consonantal phrases in which other consonantal phrases can be embedded. The depth of embedding is restricted by the notion of *strength difference*: The *matrix consonantal phrase* must be stronger (roughly, more obstruent) than the embedded one. Languages have a limit on how small a strength difference can exist between the matrix and the embedded consonantal phrase in morphologically simplex words. Such modelling correctly predicts a number of phenomena, including the phonotactic strength of affricates and the existence of emergent stops in the correct environments, e.g. *el(t)se*, *Alham(b)ra*.

Hitomi Onuma and Kuniya Nasukawa
Velar softening without precedence relations

It is generally assumed that phonological analyses, and especially segmental analyses, must refer to precedence relations between segments in order to successfully capture edge effects across boundaries and the directionality of assimilation. However, **Onuma and Nasukawa** challenge this established tradition and offer an alternative analysis of segmental phenomena without referring to precedence relations in phonological representation. As a case study, we analyse velar softening, a well-known phonological regularity in English, within the framework of Precedence-free Phonology (Nasukawa 2014, 2015, 2016; Nasukawa and Backley 2017). We propose that the process in question be analysed as an agreement effect involving the |A|-headed [|A||I|] set, which may be expressed without referring to precedence relations.

Markus Pöchtrager
Recursion and GP 2.0

Phonology is usually assumed to lack recursion. However, any such claim rests on a particular view of the workings of phonology, one that (i) countenances only a limited set of constituents that a phonological string can be broken down into and (ii) assumes that the labels of those constituents are adequate (or even relevant in the discussion).

In this article **Pöchtrager** argues, based on evidence from English vowel length, the internal structure of Putonghua vowels, Québec French vowel laxing as well as vowel reduction in general that there is reason to believe that such a view of phonology is mistaken. Work in Government Phonology (Kaye, Lowenstamm and Vergnaud 1985, 1990) has shown that mainstream concepts of constituency are often inadequate and thus questionable. By going over various such mainstream assumptions, Pöchtrager shows that in a non-arbitrary theory of phonology, recursive structures are not just convenient, but actually necessary in order to express various asymmetries we find in phonology. This is reflected in a more recent version of the theory, Government Phonology 2.0 (Pöchtrager 2006).

Clemens Poppe
Head, dependent, or both: Dependency relations in vowels

Poppe argues in favor of an element-based approach to vowel structure in which dependency between elements is defined structurally. Building on earlier work in dependency-based phonology, he proposes that the vocalic place node dominates a head place node and an optional dependent place node. Because both nodes may contain the same element, in this approach it is possible to have a three-way contrast: at the underlying level, the same element(s) may have head status, dependent status, or both. Poppe presents support for this approach to vowel structure from the vowel systems of (RP) English and Middle Korean, showing that, apart from the presence vs. absence of an element, for both languages we need to distinguish between two types of contrast: one between head and dependent elements, and one between vowels with and without identical elements. His paper concludes with a discussion of alternative dependency-based approaches, showing that, in contrast to the proposed structural approach to headedness, in these approaches it is not possible to constrain the number of identical elements in the same vowel to two.

Geoffrey Schwartz
Defining recursive entities in phonology: The Onset Prominence framework

Although it is commonly assumed that phonology is not recursive in the same way that syntax is, it is impossible to evaluate this assumption without first establishing clear definitions of the entities that are claimed not to recur. In the

Onset Prominence framework, both segments and larger prosodic constituents are derivative units that evolve from a primitive stop-vowel CV sequence. Under this view, **Schwartz** argues that each ‘segment’ is a recursion of the representational hierarchy built from the CV. Further, a recursive submersion mechanism is parametrically available, forming a range of constituents, from syllables with ‘coda’ consonants, to prosodic words and phrases. Unrestricted submersion produces configurations conducive to prosodic features traditionally associated with ‘stress-timed’ rhythm, including phonetically robust lexical stress and vowel reduction.

Ali Tifrit
Obstruent liquid clusters: Locality, projections and percolation

Tifrit aims to characterize the structure and the behavior of liquids in a slightly modified Government Phonology 2.0 framework (Pöchtrager 2006, Zivanovic and Pöchtrager 2010). He investigates the case of /Obstruent+Liquid/ (OL) clusters and proposes a representation of liquids explaining their behavior. These groups suffered misconception: the obstruent and the liquid are clearly unequally structured. The former can project while the latter cannot and, by consequence, must find a host. Tifrit discusses cases of lenition in a CV framework (Lowenstamm 1996, Scheer 2004) and illustrates the questions arising with OL clusters acting sometimes as a single element and sometimes as two distinct objects. The author underlines the theoretical issues that are related to the flatness of the CV model: *Locality* and *Infrasegmental Government*. Tifrit then reconsiders the internal content and structure of liquids. He proposes new analyses of cases of lenition, surface changes, compensatory lengthening and metatheses by formalizing them in GP2.0. Given the structures he proposes, most of the properties and behaviour of OL clusters are now expected. The main consequence of this proposal is that the problem of Locality does not arise anymore and *Infrasegmental Government* is no longer necessary.

Hisao Tokizaki
Recursive strong assignment from phonology to syntax

Tokizaki argues that stress is assigned to a constituent according to the labels assigned by the rule Set Strong. Set Strong assigns the label Strong to a set and Weak to a terminal when they are Merged. Set Strong recursively applies to syllables, words, phrases and sentences as the derivation proceeds. It is argued that stress location in derivational words is explained by Stem Stress, which is ascribed to Set Stress. In a morpheme in languages with weight-sensitive stress system, Set Stress assigns Strong to a heavy syllable, which is analyzed as a set of syllables. In a morpheme in languages with fixed stress location, Set Stress assigns Strong to a syllable that is a singleton set, which may Flip the linear order of syllables at Externalization. This analysis shows that a phonological rule Set Strong together with morphosyntactic Merge recursively applies to a set and a terminal within morphemes as well as in words and phrases, building a hierarchical prosodic structure.

Acknowledgements: I am grateful to all the contributors, not only for providing their own papers, but also for agreeing to review several other papers in the volume. In addition, the following people kindly acted as reviewers: Eugeniusz Cyran, Björn Köhnlein, Andrew Nevins, Bridget Samuels, Hidetoshi Shiraishi, Péter Szigetvári, Shanti Ulfsbjorninn, Jeroen van de Weijer. Without their valuable efforts, this volume could not have been completed. Furthermore, I thank Phillip Backley for his comments and advice on earlier versions.

This work was supported by the following MEXT/JSPS KAKENHI grants: Grant-in-Aid for Scientific Research on Innovative Areas #4903 (Evolinguistics) Grant Number JP18H05081, Grant-in-Aid for Scientific Research (S) Grant Number JP19H05589, Grant-in-Aid for Scientific Research (B) Grant Numbers JP26284067 and JP15H03213, and Grant-in-Aid for Scientific Research (C) Grant Number JP15K02611.

Kuniya Nasukawa
Vienna, June 2019

References

Chomsky, Noam. 2010. Some simple evo devo theses: how true might they be for language? In Richard Larson, Viviane Deprez and Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*, 45–62. Cambridge: Cambridge University Press.

Hauser, Marc D., Noam Chomsky and W. Tecumseh Fitch. 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1985. The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook* 2, 303–328.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7(2),193–231.

Kula, Nancy C. 2002. *The Phonology of Verbal Derivation in Bemba*. LOT Dissertation series 65. Utrecht: Holland Academic Graphics.

Kula, Nancy C. 2012. On the representation of tone in Element Theory. In Eugeniusz Cyran, Henryk Kardela and Bogdan Szymanek (eds.), *Sound Structure and Sense: Studies in Memory of Edmund Gussmann*, 353–370. Lublin: Wydawnictwo KUL.
Lowenstamm, Jean. 1996. CV as the only syllable type. In Jacques Durand and Bernard Laks (eds.), *Current trends in Phonology: Models and Methods*, 419–44. CNRS, Université Paris X and University of Salford: University of Salford Publications.
Nasukawa, Kuniya. 2014. Features and recursive structure. *Nordlyd* 41(1),1–19. *Special Issue on Features* edited by Martin Krämer, Sandra-Iulia Ronai and Peter Svenonius.
Nasukawa, Kuniya. 2015. Recursion in the lexical structure of morphemes. In Marc van Oostendorp and Henk van Riemsdijk (eds.), *Representing Structure in Phonology and Syntax*, 211–238. Berlin/Boston: Mouton de Gruyter.
Nasukawa, Kuniya. 2016. A precedence-free approach to (de-)palatalisation in Japanese. *Glossa: A Journal of General Linguistics* 1(1).9,1–21, DOI: http://dx.doi.org/10.5334/gjgl.26
Nasukawa, Kuniya and Phillip Backley. 2017. Representing moraicity in Precedence-free Phonology. *Phonological Studies* 20, 55–62.
Nespor, Marina and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
Pinker, Steven and Ray Jakendoff. 2005. The faculty of language: What's special about it? *Cognition* 95, 201–236.
Pöchtrager, Markus A. 2006. The structure of length. Ph.D. dissertation. University of Vienna.
Reuland, Eric. 2009. Language: symbolization and beyond. In Rudolf Botha and Chris Knight (eds.), *The Prehistory of Language*, 201–224. Oxford: Oxford University Press.
Samuels, Bridget. 2009. The third factor in phonology. *Biolinguistics* 3.2, 355–382.
Scheer, Tobias. 2004. *A Lateral Theory of Phonology: What Is CVCV, and Why Should It Be?* Berlin and New York: Mouton de Gruyter.
Scheer, Tobias. 2008. Why the prosodic hierarchy is a diacritic and why the interface must be direct. In Jutta M. Hartmann, Veronika Hegedüs and Henk van Riemsdijk (eds.), *Sounds of Silence: Empty Elements in Syntax and Phonology*, 145–192. Oxford: Elsevier.
Scheer, Tobias. 2011. *A Guide to Morphosyntax-phonology Interface Theories: How Extra-phonological Information Is Treated in Phonology since Trubetzkoy's Grenzsignale*. Berlin and New York: Mouton de Gruyter.
Živanovič, Sašo and Markus A. Pöchtrager. 2010. GP 2, and Putonghua, too. *Acta Linguistica Hungarica* 57(4), 357–380.

Phillip Backley and Kuniya Nasukawa

Recursion in melodic-prosodic structure

1 Introduction

There is a longstanding tradition in phonology of distinguishing between two kinds of phonological information, melodic (segmental) and prosodic (organisational). Melodic information describes individual segments, while prosodic information is concerned with the way segments are organised into larger constituents such as syllables and words. To capture this distinction between melodic and prosodic information, it is generally assumed that phonological representations consist of two independent modules, melodic structure and prosodic structure. On the one hand, melodic structure uses melodic units (e.g. features, elements) to express phonological categories and/or the phonetic characteristics of speech sounds. Meanwhile, prosodic structure uses prosodic units (e.g. nuclei, rhymes, syllables, feet, words) to specify the domains within which we find melodic units showing regular patterns. For example, sonority differences between segments are usually relevant only within the syllable domain or between adjacent syllables (Selkirk 1984, Clements 1990, Duanmu 2009), while weakening effects such as vowel reduction and consonant lenition tend to operate within the foot or word domain.

Apparently, then, there are good grounds for making a formal distinction between melodic and prosodic structure: each one employs a unique set of units, and each one encodes its own unique type of phonological information. Moreover, there is ample evidence that each one can be targeted by phonological processes independently of the other. For example, most segmental processes bring about a change in melodic structure while leaving prosodic structure (e.g. vowel/consonant length) unaffected. Conversely, vowel-glide alternations such as *i~j* and *u~w* involve a change in prosodic organisation (*i*/*u* being linked to a syllable nucleus, *j*/*w* to a syllable onset) while melodic properties remain constant.

In this paper, however, we challenge this view in which melody and prosody are kept distinct in phonological representations; instead, we propose that the two belong in a single integrated structure. A unified melodic-prosodic structure is possible, we argue, if it can be shown that the same units are able to represent both melodic (segmental) and prosodic (organising) properties. Below we demonstrate how phonological elements (Harris and Lindsey 1995, Backley 2011), which are conventionally used to represent only melodic properties, may also take on an organising role and be used in place of standard

https://doi.org/10.1515/9781501512582-002

prosodic constituents. Furthermore, we show how the use of recursive structure is integral to the well-formedness of the unified (melody-prosody) model being proposed.

2 Melody-prosody integration

2.1 Motivation

Although most scholars still adhere to the mainstream view that melody and prosody should be separated in representations, there are also arguments for *not* representing melody and prosody as independent entities. This is the position we defend here, our motivation resting on two factors which suggest the need for a unified melodic-prosodic structure.

First, melody and prosody should be integrated because the two regularly interact. Clearly, interaction is possible even between autonomous parts of a representation, but if these different parts refer to different structural units, then we are forced to conclude that any such interaction is based on random relations. For example, in languages such as English and Swedish, aspirated stops are usually restricted to syllable onsets. That is, the melodic unit which represents stop aspiration – the element |$\underline{H}$| (or alternatively, the feature [constricted glottis]) – invariably appears in a syllable onset rather than a nucleus or a coda. However, it is not obvious how a formal link between the melodic unit |$\underline{H}$| and the prosodic unit 'onset' can be expressed. These two units belong to different vocabularies, making the relation between them no more than a stipulation. But if melody and prosody are unified into a single structure and represented in terms of the same units, then it may be possible to explain – rather than merely describe – why a given melodic property tends to be associated with a given prosodic property.

A second reason for rejecting the traditional division between melody and prosody is linked to the idea of empty structure. Government Phonology (e.g. Harris 1990; Kaye, Lowenstamm and Vergnaud 1990; Cyran 2010; Charette 1991) and its offshoots (e.g. van der Hulst 2003, Scheer 2004) employ representations in which a prosodic unit can be pronounced even if it has no melodic units associated with it. The typical case is a melodically unspecified or 'empty' nucleus, which may be phonetically realised as a default vowel such as [ə] or [ɨ] if the required prosodic conditions are met. In a standard feature-based approach to segmental structure, this would be considered an anomaly because, if features are responsible for defining a segment's phonetic qualities, then the absence of features should equate to silence.

But in a Government Phonology approach, where segments are represented by elements rather than by features, the same outcome is legitimate. This is because elements encode marked phonological properties rather than phonetic qualities. So, the absence of elements merely expresses the absence of marked properties – which means that the empty nucleus can still be realised as an unmarked or default vowel. In this way, an audible segment can be pronounced in a position which contains no melodic units. This has the effect of blurring the division between melody and prosody. In traditional terms, melodic structure represents segmental information while prosodic structure represents relational or organising information; but in the Government Phonology approach just described, this distinction breaks down as we find segments being associated with prosodic rather than melodic structure. This provides a further reason for rejecting the standard melody-prosody distinction, and instead, for combining the two into a unified representation.

If melodic and prosodic information are to be integrated into a single structure, then it makes sense for both to be 'speaking the same language' by using the same structural units. Following the Precedence-free Phonology approach described in Nasukawa and Backley (2015), we propose to eliminate from representations the conventional labels for prosodic constituents (e.g. onset, nucleus, syllable, foot, word) and replace them with elements – the units which, until now, have been associated only with melodic structure. In employing the same units at all levels of representation, we move closer towards our goal of unifying melodic and prosodic structure by avoiding the need to refer to units which specify only one kind of information (i.e. melodic or prosodic). Our claim, therefore, is that phonological representations refer only to elements: the elements in melodic structure have an interpretive function and provide information about segmental properties, while the elements in prosodic structure take on an organising function and provide information about relational or organizing properties. After all, phonological representations are primarily concerned with segmental expressions and how these are organized in morphemes. And because morphemes are identified by their melodic properties, it follows that they should be represented using only the units of melodic structure, i.e. elements.[1]

1 Alternative approaches, in which all melodic properties are represented in terms of structural (organising) properties, are discussed elsewhere in this volume. In particular, the reader is referred to the contributions by Edoardo Cavirani and Marc van Oostendorp, Markus Pöchtrager, Geoff Schwartz, and Marcel den Dikken and Harry G. van der Hulst.

Below we show how elements are phonetically interpreted at the melodic level, while at the prosodic level they enter into head-dependency relations with one another. We argue that these head-dependency relations account for the phonotactic and distributional patterns which we observe in morphemes and which are traditionally expressed in terms of prosodic structure.

2.2 Rethinking hierarchical structure

To reiterate the main point, we propose that elements function not only as melodic units but also as prosodic constituents. However, it emerges that not all elements behave this way – it is chiefly the resonance elements |A|, |I| and |U| that have this dual function. This is not surprising, given that these elements are primarily associated with nuclei, and that it is nuclei which function as the building blocks of prosodic structure (cf. onsets, which are mostly irrelevant to higher-level prosodic relations). In traditional descriptions of hierarchical prosodic structure, a nucleus first projects to a rhyme node, then to a syllable, then to a foot, and so on. The question, then, is how this familiar representation of the prosodic hierarchy will change if we pursue an approach in which melody and prosody are unified into a single structure.

If there is no division between melodic structure and prosodic structure, then logically, there is no interface between the two. This state-of-affairs marks a clear departure from the traditional approach to phonological structure, in which it is assumed (i) that the lowest level of the prosodic hierarchy consists of terminal units – either syllabic constituents such as nuclei or bare timing slots such as skeletal positions – and (ii) that these terminal units interface with units of melodic structure such as features or elements. But by pursuing the idea being proposed here, that melody and prosody form a single structure, we are forced to abandon the assumption that prosodic structure terminates at the point where it meets the melodic (segmental) level. Instead, prosody and melody may be viewed as contiguous parts of one continuous hierarchy.

In a hierarchical model, a structural head has scope over everything it dominates. For example, in a standard view of prosodic structure the head of a foot 'contains' all the material associated with the nuclei immediately below it. Putting this another way, the foot node is the instantiation of its constituent properties – that is, it embodies the properties specified in its dependent syllables. And a similar relation holds between all adjacent levels on the hierarchy. In the case of a nucleus, which is usually regarded as a terminal node on the prosodic hierarchy, it may instead be viewed as the instantiation of all the vocalic properties that are associated with it. These properties may be expressed by features such as [±high] and [±back] or, following the Element Theory approach we employ here, by the resonance elements |I|, |U| and |A|. In other words, the unit conventionally labelled 'nucleus' is nothing more than the embodiment of its constituent elements. We argue that these melodic elements are associated directly with nuclei, so there is no need to posit any intervening level of structure (e.g. timing slots) or to refer to any interface between melody and prosody.

The details of this unified melodic-prosodic hierarchy will be described in section 4. This is preceded by an overview of element representations. We describe how elements differ from standard features in some fundamental ways, making them ideally suited to the recursion-based hierarchical model being developed here.

3 Element-based vowel representations

3.1 Elements

Like features, elements are units of melodic structure which represent phonological categories. Unlike features, however, they are associated with acoustic patterns in the speech signal rather than with properties of articulation (Harris and Lindsey 1995, Nasukawa and Backley 2008, Backley 2011, Nasukawa 2017). The relevant acoustic patterns are those that are thought to be linguistically significant – that is, they carry linguistic information about the identity of morphemes. For example, the element |H| represents the pattern of aperiodic noise energy that is observed in fricatives and in the release phase of stops, while the element |U| represents a formant pattern in which sound energy is concentrated at the lower end of the spectrum, as is found in labials, velars and rounded vowels (Nasukawa and Backley 2008, Backley and Nasukawa 2009, Backley 2011).

Although Element Theory exists in several forms (Backley 2012), standard versions use the six elements shown in (1). Each element is associated with its own unique acoustic pattern (Harris and Lindsey 1995; Nasukawa and Backley 2008, 2011; Backley and Nasukawa 2009, 2010).

(1) Elements and their acoustic patterns

a. Vowel (resonance) elements

|I|	'dip'	low F1 with high spectral peak – convergence of F2 and F3
|U|	'rump'	low spectral peak – lowering of all formants
|A|	'mass'	central spectral energy mass – convergence of F1 and F2

b. Consonant (laryngeal) elements

|ʔ|	'edge'	abrupt and sustained drop in amplitude
|H|	'noise'	aperiodicity, noise
|L|	'murmur'	periodicity, nasal murmur

As shown here, the elements naturally divide into two subsets. The resonance elements |A|, |I| and |U| are associated with patterns which relate to formant structure, so they are primarily associated with vowels. Meanwhile, the laryngeal elements |H|, |L| and |ʔ| refer to other properties of the speech signal such as noise energy and amplitude, so they appear mainly in the representation of consonants.

Elements do not just refer to aspects of the physical speech signal, however. They are also linked to the abstract phonological categories that are present in mental representations. These representations are used by native speakers to identify individual morphemes and words. Furthermore, each of the acoustic cues described in (1) is directly associated with a particular phonological category. Note that these linguistic categories do not always respect the traditional division between vowels and consonants. For example, the formant cues associated with the 'vowel' elements |I|, |U| and |A| describe vowel quality; but in addition, they distinguish consonant place properties too. So, at an abstract level the 'vowel' elements contribute to the representation of both vowels and consonants. Similarly, the 'consonant' elements |ʔ|, |H| and |L| capture the characteristics of consonants such as the presence of noise energy and rapid changes in amplitude; but they also refer to vowel properties which are contrastive in some vowel systems, such as nasality and lexical tone. The following table illustrates how each element contributes to nuclear and non-nuclear expressions.

(2) Elements and their phonological properties

a. Vowel (resonance) elements

	nuclear	*non-nuclear*
|I|	front vowels	coronal: dental, palatal POA
|U|	rounded vowels	dorsal: labial, velar POA
|A|	non-high vowels	guttural: uvular, pharyngeal POA

b. Consonant (laryngeal) elements

	non-nuclear	*nuclear*
|ʔ|	oral/glottal occlusion	creaky voice (laryngeal vowels)
|H|	aspiration, voicelessness	high tone
|L|	nasality, obstruent voicing	nasality, low tone

Whereas orthodox distinctive features are bivalent (i.e. they have a plus and a minus value), elements are monovalent or single-valued. This means that lexical contrasts are expressed in terms of an element's presence/absence rather than in terms of its plus/minus value. Another characteristic of elements is their 'autonomous interpretation', which allows an element to be phonetically interpreted by itself since it has its own 'autonomous phonetic signature' (Harris and Lindsey 1995: 34). Having said that, expressions usually involve a combination of elements. In standard Element Theory (Harris 1994, Backley 2011), for example, the elements |I| and |A| are realised individually as [i] and [ɑ] respectively. But they can also combine to form a complex expression |I A|, which is pronounced as a front mid vowel such as [e] or [ɛ]. Consonants are also represented by complex expressions. For example, |U| (labiality), |ʔ| (occlusion) and |H| (noise) are realised individually as [w], [ʔ] and [h] respectively, whereas in combination they are pronounced as [p] – the realisation of the complex expression |U ʔ H|.

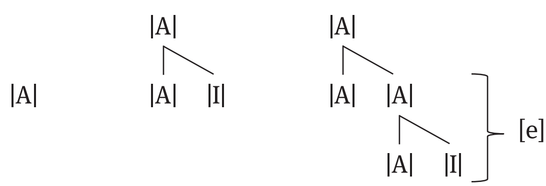
3.2 |I|, |U|, |A| as prosodic constituents

In phonological representation, prosodic structure is normally based on relations between rhymes/nuclei – onsets are rarely involved in prosodic patterning. And this is reflected in the Precedence-free Phonology model of representation that we employ here, in which it is exclusively the vowel elements |I|, |U| and |A| – the units that encode the contrastive properties of nuclei – which have a prosodic function. We argue that |I|, |U| and |A| not only carry lexical information about the identity of vowel segments, they also project beyond the melodic structure to higher prosodic levels, where they form head-dependent relations with one another. As these asymmetric relations progress upwards through the prosodic hierarchy, they mark out a series of successively wider prosodic domains which correspond to traditional units such as rhyme, syllable, foot and word. Because this can be achieved by referring only to |I|, |U| and |A|, there is no need to introduce the constituent labels 'rhyme', 'syllable', 'foot' and 'word' into the structure. The result is a representation which integrates melodic and prosodic information into a unified melody-prosody structure, but one which minimizes the number of different structural units it uses. Even 'nucleus' is not recognised as a formal constituent, since a nucleus is nothing more than an instantiation of the melodic units (i.e. |I|, |U| and |A|) associated with it. (Note that, for convenience, we will continue to use the term 'nucleus' as an informal label for the prosodic domain associated with a vowel. Strictly speaking, however, there is no prosodic constituent called 'nucleus' in the representations proposed here.)

Like all constituents in this integrated melodic-prosodic structure, a nucleus contains units (elements) that are combined via head-dependent relations. The head of a nucleus must be one of the vowel elements |I|, |U| or |A|, the choice being language-specific and determined by parameter. In English, for example, the head of a nucleus is |A|. This head element can either stand alone, or it can support a dependent element also from the |I|/|U|/|A| set. If a head element stands alone, it is pronounced as a weak or default vowel, as in (3a) (the representation of weak vowels will be discussed in section 3.3). But if it takes a dependent, then the whole expression is realized phonetically as a full or lexically contrastive vowel, as in (3bc) (the representation of full vowels will be discussed in section 3.4). Below it will be shown how the quality of a full vowel derives largely from the properties of its dependent element(s), rather than from those of its head.

(3) |I|, |U|, |A| as prosodic constituents

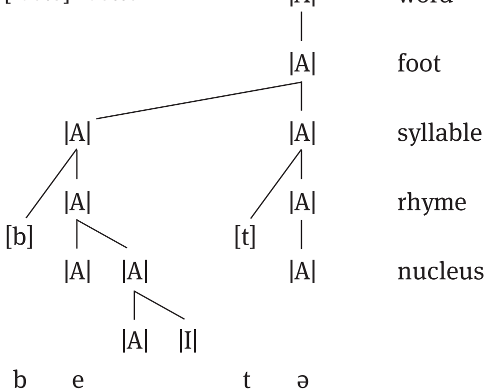
a. [ə] b. [i] c. [e]



In this model, phonological structure is assumed to be recursive. That is, a dependent element can have a dependent of its own, which will occupy a lower (i.e. more deeply embedded) position in the structure. For example, the mid vowel in (3c) requires one more level of embedding than the high vowel in (3b). In principle, there is no restriction on the amount of complexity (embedding) that a structure may have. But on the other hand, representations are never more complex than they need to be – they must be complex enough to express the set of contrasts in a language, but that is all. So, the complexity of a language's vowel system will always dictate the number of levels of embedding required in representations for a given language. In all cases, vowel structure involves chains of binary head-dependency relations holding between tokens of the vowel elements |I|, |U| and |A|.

In addition to supporting dependent structure, the head of a nucleus also projects upwards in its role as a prosodic head. And depending on the prosodic level in question, the head element can function as the head of a syllable-sized domain, a foot-sized domain, or a word-sized domain. This is illustrated in (4), where consonants are shown as whole segments since consonant structure is not relevant to the issue of head projection.

(4) [ˈbetə] ‘better’



The structure of consonants will be described in more detail in section 4. For the moment, it may be noted that (4) departs from the conventions of Dependency Phonology (Anderson and Ewen 1987) and Government Phonology (Harris and Kaye 1990, Harris 1994) by representing trochaic words as right-headed structures. This follows Nasukawa and Backley (2015: 68), where it is claimed that constituent heads are important structurally but have a low informational load, while dependents are less important for structure-building but are rich in terms of information. For an overview of the Precedence-free Phonology approach, and a detailed discussion of the relation between phonological structure and phonetic realisation, see Nasukawa (2017).

If elements can function as prosodic units in this way, then phonological structures need only refer to elements and to the head-dependency relations holding between them – traditional prosodic labelling (‘rhyme’, ‘syllable’, etc.) becomes superfluous. As (4) shows, the head element projects upwards to every level of the prosodic hierarchy, defining successively wider prosodic domains as it does so. In this way, it is still possible to identify prosodic domains for the purposes of describing phonological patterns, but it can be done without referring to the usual constituent labels.

There are at least two advantages of adopting this approach. First, we avoid having to use constituents that are specific to just one level of structure – for example, ‘syllable’ only refers to the syllable level; and at the same time, we minimize the inventory of structural units employed in representations. Second, we make it easier to understand how and why melody and prosody interact – that is, why melodic (segmental) patterns are often sensitive to their prosodic context. For example, vowel reduction in English occurs in the weak part of a foot, but when this pattern is described in traditional terms it needs to be stipulated, as the relation between the prosodic label 'foot' and the melodic units |I|/|U|/|A| appears to be arbitrary. On the other hand, if the same units are used to describe both melodic and prosodic structure, then melody-prosody interaction begins to 'make sense' as a potential or even expected way for languages to behave.

So, the Precedence-free Phonology model being described here reinforces the idea of a unified representation in which elements are used to represent phonological information at every level of melodic and prosodic structure. These elements fulfil their familiar role as interpretable units of melody, but they also have an organising function by concatenating recursively via head-dependency relations to create successively larger prosodic domains – and thus, to generate successively larger phonological strings.

3.3 Empty nuclei and default vowels

The motivation for allowing only |I|, |U| and |A| to function as prosodic heads comes from the way that so-called 'empty' nuclei are phonetically realized. In the previous section we argued for a unified representation in which the traditional split between melodic structure and prosodic structure is obscured. And this view is supported within the government/licensing approach to representation (Charette 1991, 2003; Harris 1997; Kaye 2000; Cyran 2010; Scheer 2004), in which the distinction between melody and prosody has always been somewhat blurred. For example, a nucleus may be pronounced even when it is empty – that is, when it has no elements associated with it. In other words, it is possible for prosodic structure to be phonetically realized even in the absence of lexical melodic structure.

An empty nucleus typically functions as a default vowel and is pronounced with a central or non-peripheral quality somewhere within the range [ə]~[ɨ]~[ɯ]. Default vowels often appear in loanwords, when the native phonology requires a nucleus to be pronounced and there is no lexical vowel in the original borrowed form. For example, English has a schwa-like vowel between consonants that cannot form a complex onset, e.g. [gəˈdænsk] *Gdansk*, [kəˈnjuːt] *Cnut* (or *Canute*). In Japanese the quality of a default vowel is closer to a high back [ɯ], which breaks up consonant sequences in loanwords such as [takɯɕiː] *taxi*, [akɯserɯ] *accel(erator)*. Because the precise quality of a default vowel varies from one language to another, it is usually treated in terms of a language-specific parameter.

Vowel qualities such as [ə]~[ɨ]~[ɯ] make ideal default vowels because they are usually non-contrastive in vowel systems; after all, the purpose of a default vowel is to fill a nucleus without introducing any new linguistic (contrastive) information. In languages showing vowel reduction effects, such as English, these are also the vowel qualities that occupy weak prosodic positions. Because default vowels ([ə]~[ɨ]~[ɯ]) are associated with weak syllables, and because their distribution (in weak syllables) is complementary to that of full vowels (in strong syllables), we will claim that [ə], [ɨ] and [ɯ] are the weak realizations of the three full vowels that are phonetically closest to them – namely, [a], [i] and [u] respectively (Nasukawa 2014). Furthermore, because the full vowels [a], [i] and [u] are represented by the resonance elements |A|, |I| and |U|, we will assume that the same elements |A|, |I| and |U| are also latently present in their weak counterparts [ə], [ɨ] and [ɯ] – that is, in so-called 'empty' nuclei. The structures for [ə], [ɨ] and [ɯ] are given in (5). Note that these vowels are represented by 'minimal' structures: when a lone |A|, |I| or |U| stands as a single-element expression with no dependent structure, it is realised as a central vowel [ə], [ɨ] or [ɯ].

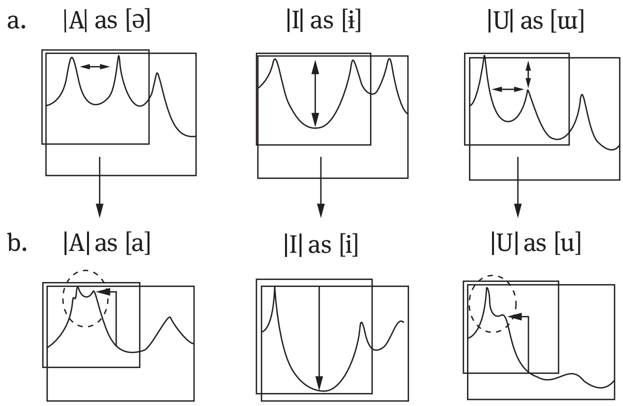
(5) |I|, |U|, |A| as default vowels

<table>
<tr><td>a. [ə] (English)</td><td>b. [ɨ] (Cilungu)</td><td>c. [ɯ] (Japanese)</td><td></td></tr>
<tr><td>|A|″</td><td>|I|″</td><td>|U|″</td><td>'syllable'</td></tr>
<tr><td>|A|′</td><td>|I|′</td><td>|U|′</td><td>'rhyme'</td></tr>
<tr><td>|A|</td><td>|I|</td><td>|U|</td><td>'nucleus'</td></tr>
</table>

These structures represent vowels in their most basic form: they are weak, non-peripheral vowels because they contain just a single element. And because they have no dependent structure, they carry no contrastive/lexical information. In this sense, the element that is present in each expression has a purely structural role: it functions as a prosodic constituent (i.e. a nucleus) and it can also be phonetically interpreted as a default (i.e. non-lexical) vowel. We observe a typological split between languages based on the quality of their default vowel. As shown in (5), languages with |A| as their head element have a schwa-like default vowel (e.g. English), those with |I| have a high central [ɨ] (e.g. Cilungu), and those with [U] have a back [ɯ] as their default vowel (e.g. Japanese). In acoustic terms, a latent element provides the phonetic baseline onto which other elements' acoustic patterns are superimposed. However, if no other elements are present (i.e. in an empty nucleus) then this baseline resonance is exposed and the head element becomes audible.

Support for the representations in (5) comes from physical evidence relating to the acoustic properties of the vowels in question. Unlike distinctive features, which mostly refer to properties of articulation (e.g. tongue position in [high] and [back], lip shape in [round]), elements are associated with acoustic patterns in the speech signal. We should therefore expect to find a similar acoustic shape in vowels that contain the same element. Specifically, the vowels in each of the weak-strong pairs [ə]-[a], [ɨ]-[i] and [ɯ]-[u] ought to have spectral patterns that are, at least to some extent, alike (Nasukawa 2014).

(6) Default vowels and full vowels compared



Consider first the spectral shape of [a] in (6b), which appears to be an exaggerated version of the equivalent shape for the corresponding weak vowel [ə]. The ‘mass’ pattern (see figure (1) above) associated with the element |A| is characterised by F1-F2 convergence, and this pattern is more prominent in [a] than in [ə] since the F1 and F2 energy peaks are closer together (i.e. they fully converge). Turning to strong [i] versus weak [ɨ], the difference again comes down to the prominence or salience of the relevant acoustic pattern. The ‘dip’ pattern associated with |I| is marked by a high F2 peak, which creates a trough or dip between F1 and F2. The trough in [i] is visibly deeper and more prominent than in [ɨ], and for this reason it may be understood as an exaggerated form of ‘dip’. Finally, [u] and [ɯ] both display the ‘rump’ pattern associated with |U|, in which acoustic energy is concentrated at low frequencies. This produces a falling spectral shape which is sharper and more exaggerated in strong [u] than in weak [ɯ].

The acoustic similarities shown in (6) lend support to the idea that a so-called ‘empty’ nucleus is in fact not empty, because it contains a latent element which is realized as baseline resonance. And this relates to the point made earlier concerning the labelling of prosodic units: if an ‘empty’ or unspecified nucleus contains a default element, then this default element is enough to represent the nucleus in question – there is no need for an additional constituent label ‘nucleus’ because the nucleus is already defined by a default |I|, |U| or |A|. The same applies to constituents at higher prosodic levels too. As illustrated by the structure in (4), the element which serves as the head of a nucleus is projected to successively higher levels and becomes the head of successively larger prosodic domains. At each level this head element defines the prosodic domain in question, so there is no advantage in renaming these domains using arbitrary labels such as ‘syllable’, ‘foot’ and ‘word’.

3.4 Contrastive vowels

The minimal structures in (5) are pronounced as weak vowels rather than as full vowels because a lone head element produces only baseline or default resonance. That is, a minimal structure can express only a minimal amount of phonological information: it signals the presence of a nucleus, but one which has no lexical/contrastive properties. By contrast, a full vowel has a more complex structure containing dependent elements, and it is this additional structure which expresses contrastive properties. As shown in (7b) and (7c), an endocentric head-dependency relation between elements increases structural complexity, which in turn allows an expression to carry lexical information.

(7) Endocentric head-dependency

a. [ə]

|A|″
|A|′
|A|

b. [a]

|A|″
|A|′
|A| |A|

c. [i]

|A|″
|A|′
|A| |I|

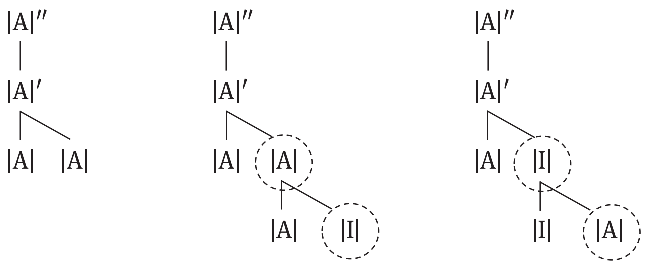
In (7a), repeated from (5), the lone head element |A| is pronounced as a schwa-like vowel; this is the realization of (non-contrastive) baseline resonance. But in (7b) this head |A| takes another token of |A| as a dependent, and the ‘mass’ (high F1) acoustic pattern of this dependent |A| is superimposed onto the schwa-type baseline. In effect, the acoustic pattern associated with the dependent element masks the baseline resonance and listeners perceive a low ‘mass’ vowel [a]. In (7c) too, the baseline resonance is inaudible because it is overridden by the properties of a dependent element; the dependent |I| means that the expression is realised as [i].

Using the simple structures in (7) it is possible to represent three contrastive vowels: [i] (dependent |I|), [u] (dependent |U|) and [a] (dependent |A|). But clearly, this is not enough for describing the vowel systems of most languages. To express additional vowel contrasts we need to allow for further element combinations, which means introducing more levels of embedding into the structure. Note that, although element embedding is a characteristic of the approach being developed here, it follows the conventions of element-based phonology by requiring elements to combine asymmetrically. In standard versions of Element Theory (Harris 1994, Cyran 1997, Backley 2011), as well as in Dependency Phonology (Anderson and Jones 1974, Anderson and Ewen 1987) and Particle Phonology (Schane 1984, 1995), mid vowels are represented by element compounds in which the relative salience of heads and dependents affects phonetic realization. For example, the expressions $|\underline{\mathrm{I}}\ \mathrm{A}|$ and $|\mathrm{I}\ \underline{\mathrm{A}}|$ contain the same elements but the difference in their headedness makes them phonetically distinct: the $|\underline{\mathrm{I}}|$-headed structure $|\underline{\mathrm{I}}\ \mathrm{A}|$ is realized as [e] while $|\underline{\mathrm{A}}|$-headed $|\mathrm{I}\ \underline{\mathrm{A}}|$ is pronounced as a more open [æ].

The recursive model of melodic representation being developed here also requires an asymmetry between elements, but it expresses this relation structurally rather than by using a diacritic (i.e. an underline). This is illustrated by the structures in (8).

(8) Vowels with complex structures

a. [a] b. [e] c. [æ]



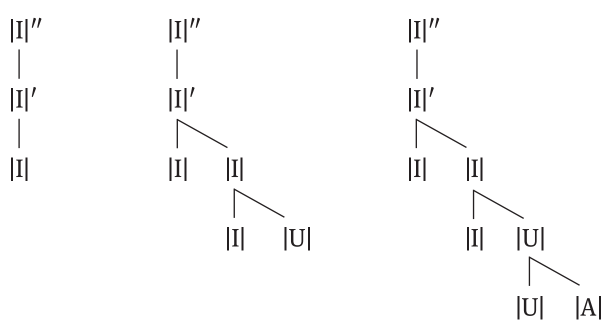
It is the most deeply embedded part of the structure – that is, the 'lowest' dependent – which makes the biggest contribution to the phonetic interpretation of an expression. In (8a) dependent |A| is the only unit in the structure which carries contrastive information, so the expression is realized as a low vowel [a]. (8b) contains an additional level of embedding and therefore two dependent elements (circled). Here the palatal resonance associated with dependent |I| predominates over its local head element |A| to produce a high mid [e].

In (8c) this asymmetric relation between |A| and |I| is reversed, with |A| in the most embedded part of the structure; since the ‘mass’ properties of this dependent |A| are more prominent than the ‘dip’ properties of its local head element |I|, the result is a more open [æ]. In all cases, then, phonetic interpretation depends not only on which elements are present, but also on the position of each element in the hierarchy of head-dependent relations.

In principle, there is no limit to the number of levels of embedding that a structure may have. But at the same time, representations are never more complex than they need to be. The grammar of a language must generate a set of melodic structures which is big enough to capture all the lexical contrasts in that language, but no more. And in most languages (i.e. those with more than three contrastive vowels) this will require element concatenation. Every instance of element concatenation introduces a new head-dependency relation, and therefore, an additional level of embedding. In a typical triangular vowel system comprising [a i u e o], only the structures in (8a) (for [a i u]) and (8b) (for [e o]) are needed. Compare this with a language such as Turkish, which requires an extra level of embedding to accommodate additional vowels such as [ü] and [ö], as in (9bc). Note that the structural head in Turkish is |I| rather than |A|, which is reflected in the [ɨ] quality of its baseline resonance – see (5) above.

(9) Vowel structures in Turkish

a. [ɨ] b. [ü] c. [ö]



As these examples demonstrate, successive levels of embedding are introduced in a recursive fashion until all the required vowel contrasts are uniquely represented. Note that there is a direct relation between the complexity of an expression and the complexity of the resulting speech signal when that expression is pronounced, since each dependent element produces a unique modulation of the carrier signal away from its baseline pattern (Harris 2005, 2009).

What emerges from this discussion is that, in the Precedence-free Phonology approach, heads and dependents have quite different roles in phonological structure (Nasukawa and Backley 2015). Head elements are important for structure-building because (i) they support dependent elements and (ii) they project to higher prosodic levels. On the other hand, they are not important for phonetic realization: in full vowels the head element is masked by the acoustic properties of its dependent(s); it is only in the absence of dependent structure that the head element is heard – and even then it is realized as baseline resonance, which carries no melodic information (Nasukawa 2014, 2016, 2017; Nasukawa and Backley 2005). Meanwhile, the opposite is true for dependents: they are unimportant for structure-building because they are merely added to existing structure and do not project to higher prosodic levels; but they do make an important contribution to phonetic interpretation because they represent the most salient melodic properties in a complex expression.

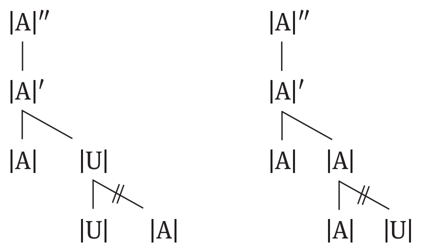
3.5 Vowel weakening

Models of vowel representation must express the lexical contrasts and natural classes that are observed across languages. They should also capture aspects of dynamic behaviour such as vowel weakening and other dynamic phonological effects. Crosswhite (2000) and others describe two kinds of vowel weakening motivated by two different forces: centrifugal systems are driven by contrast enhancement, which neutralizes contrasts in favour of the peripheral vowels [a], [i] and [u], while centripetal systems aim for prominence reduction and produce reduced vowels with a central quality such as [ə] or [ɨ]. Harris (2005), on the other hand, develops a unified approach in which all instances of vowel weakening derive from the same mechanism – namely, the suppression of dependent element structure in weak positions. We adopt Harris's approach here and apply it to the hierarchical element structures described above.

In Element Theory it is assumed that after part of an element expression is suppressed, speakers can still pronounce any remaining parts of the structure. In some cases, this reduced structure will still be complex, i.e. its head element will have at least one dependent, as in (10ab). In other cases, it will lose all its dependents to leave a minimal structure consisting of just a bare head element, as in (10c).

(10) Element suppression

a. [o] > [u] b. [o] > [a] c. [a] > [ə]



The vowel reduction pattern in (10a) is found in Bulgarian, where stressed [e o] alternate with [i u] in unstressed syllables, e.g. *r*[ó]*guf* 'of horn', *r*[u]*gât* 'horned' (Petterson and Wood 1987). This change involves suppressing the element that is lowest in the structure, namely, dependent |A|. When this happens, the |U| immediately above dependent |A| remains intact; and because this |U| is also a dependent (of the ultimate head |A|), it contributes to phonetic interpretation. By itself, dependent |U| is realized as [u]. Meanwhile, the weakening effect in (10b) is observed in Russian (Crosswhite 2000: 110); and again, the most deeply embedded element (dependent |U|) is suppressed, leaving behind its local head (the |A| above it). Because this |A| is also a dependent (again, of the ultimate head |A|), it is pronounced. On its own, dependent |A| is realized as [a].[2] In (10c) too, the lowest element in the structure is targeted. In this case, however, it leaves behind only the head element. Recall from (7a) that the ultimate head of an expression functions as a default vowel and is phonetically realized as baseline resonance (here, a weak [ə]).

This approach to vowel weakening makes two assumptions (Backley and Nasukawa 2018). First, it assumes that vowel reduction operates blindly and uniformly – the process always targets the most deeply embedded layer(s) of a vowel's structure. Second, reduction is a structure-depleting process, meaning that an expression which undergoes weakening always loses some of its structural complexity. And these two assumptions lead to some interesting observations about the abstractness of element-based representations – and indeed, about the abstractness of phonetic symbols. Consider, for example, the following patterns of mid vowel reduction (neutralisation) in Italian and Slovene.

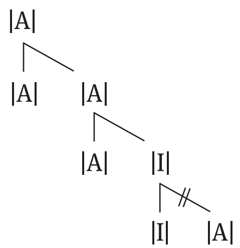
2 There is no anomaly in the fact that the dependency relation between |U| and |A| is different in Bulgarian [o] (10a) and Russian [o] (10b). Element structures are primarily a reflection of phonological rather than phonetic properties, and consequently, phonetically similar sounds can have non-identical representations if they function differently in different languages.

(11) Vowel reduction in Italian

a. Italian (Krämer 2009: 100)

	stressed	*unstressed*
[ɛ] > [e]	[ortoˈpɛdiko] ‘orthopaedist’	[ortopeˈdiːa] ‘orthopaedics’
[ɔ] > [o]	[ˈlɔʤika] ‘logics’	[loʤikaˈmente] ‘logically’

b.

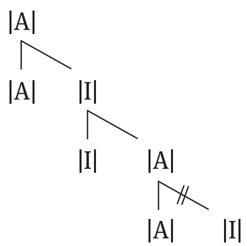


(12) Vowel reduction in Slovene

a. Slovene (Bidwell 1969, Crosswhite 2001: 31)

	stressed	*unstressed*
[e] > [ɛ]	[ˈr**e**ːʧ] ‘word’ nom. sg.	[r**ɛ**ˈʧiː] ‘word’ gen. sg.
[o] > [ɔ]	[ˈm**o**ʒ] ‘man’ nom. sg.	[ˈm**ɔ**ˈʒjeː] ‘men’ nom. pl.

b.



The four mid vowels [e ɛ o ɔ] are present in Italian and Slovene, pronounced with similar phonetic qualities in both languages. However, vowel weakening reveals that these vowels have different phonological identities in the two systems. The alternations in (11a) suggest that in Italian [e o] must be structurally less complex than [ɛ ɔ], since vowel reduction is a structure-depleting process and it is [e o] which appear in weak (unstressed) positions. But in Slovene the opposite is true – in (12a) [e o] weaken to [ɛ ɔ], so [ɛ ɔ] must be less complex than their tense counterparts [e o]. This difference does not derive from the vowel reduction process itself, which operates uniformly by suppressing the lowest element(s) in any target structure. Rather, it must result from the way elements are organised in the target structures concerned, as can be seen by comparing (11b) with (12b): in Italian the full vowel [ɛ] in (11b) weakens to [e] when its lowest dependent |A| is suppressed, while in Slovene the full vowel [e] in (12b) reduces to [ɛ] by losing dependent |I|.

This difference suggests that the relation between phonological structure and phonetic realisation is an indirect one – something that Element Theory has always maintained. Since element expressions are mental objects, there is no precise or consistent correspondence between elements and the physical (e.g. articulatory) properties of spoken language. So, to determine a vowel's element structure we focus primarily on its phonological behaviour rather than on its phonetic properties. This point is highlighted in the above examples, which illustrate how processes such as vowel weakening can shed light on phonological representations, regardless of their precise phonetic qualities. If weakening operates blindly on any target vowel, then the typological differences we find – between Italian 'tensing' in (11) and Slovene 'laxing' in (12), for instance, and between centrifugal and centripetal vowel reduction systems – cannot be accounted for by assuming that different structure-changing mechanisms are at work. Instead, they must stem from differences in the way individual vowels are represented in terms of their hierarchical element structures.

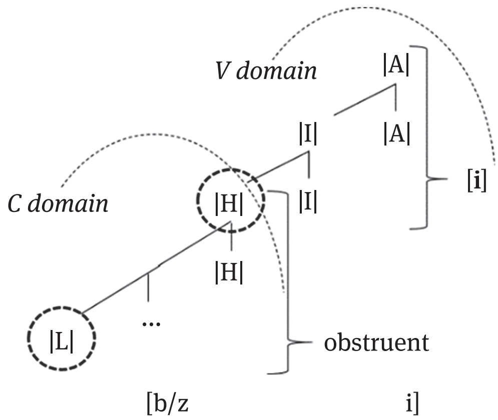
4 Recursion in consonant structure

It was mentioned above that, although elements can have a prosodic function in addition to their melodic function, this only applies to the resonance elements |I|, |U|, |A|. The remaining elements – the non-resonance elements |H|, |L|, |ʔ| – serve a more conventional role: they represent segmental categories in consonants. In this section we show how element structure in consonants, like that in vowels, is recursive; that is, element concatenation again makes use of successive layers of element embedding.

In all languages, C and V combine to form a basic prosodic unit in which the two constituents have unequal status: C is dependent on V. This is captured in syllable structure terms by saying that a rhyme (containing V) takes an onset (containing C) as its dependent. And although the representational approach described in section 3 rejects conventional notions of syllable structure, it can still express the same asymmetric relation. It does this by positioning consonant expressions below vowel expressions on a single hierarchical element structure. A syllable-sized CV unit such as [bi] or [zi] (an obstruent followed by a high front vowel) is thus represented as in (13).[3]

3 As the name implies, Precedence-free Phonology uses representations in which no reference is made to any precedence relations between sounds. In terms of linearity, therefore, there is no difference between right-branching and left-branching (Nasukawa 2011).

(13) A syllable-sized unit (Precedence-free Phonology model)



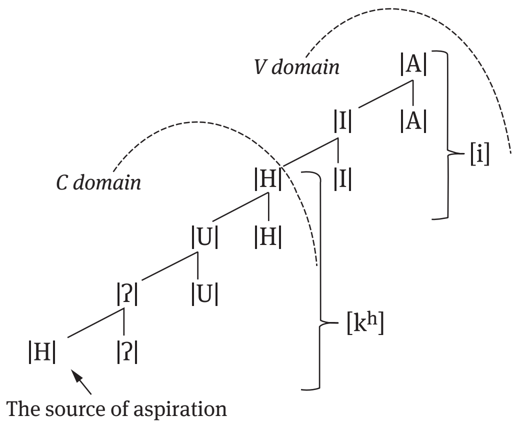
The V-domain in the upper part of (13) contains the element structure for the vowel [i]. It has |A| as its structural head, which takes |I| as a dependent. And since dependents make a bigger contribution to phonetic realization than heads, the entire V-domain is pronounced as [i] (see (7c) above). Then the C-domain, as a dependent of the V-domain, is embedded within it. The head of the C-domain is |H|, which is a dependent of the lowest element in the V-domain, namely |I|. The appearance of the noise element |H| indicates that the structure from this point downwards has the characteristics of a consonant. Consonant structure is built up in the same way as vowel structure, with additional elements being concatenated by introducing further levels of embedding. This will be illustrated in (14) below.

Representing a V-domain and a C-domain as a unified structure reflects the fact that, in phonological terms, the two behave as a single, syllable-sized prosodic unit. Nevertheless, in phonetic terms each domain is distinct – we perceive a consonant sound followed by a vowel sound. This derives from the fact that the upper and lower parts of the unified CV structure have incompatible phonetic (physiological) properties: the upper domain is vocalic while the lower domain is consonantal. And as such, they cannot be realised simultaneously;[4] for speakers, the only option is to pronounce them in sequence. The question, however, is how language users determine the order of C and V sounds in a sequence, if this information is not encoded explicitly in representations. Using the CV structure in (14) we illustrate how the linear ordering of

4 Languages do not have, for example, obstruent vowels or vocalic obstruents (Ladefoged and Maddieson 1996).

individual sounds falls out from the network of head-dependent relations in an expression (Nasukawa, Backley, Yasugi and Koizumi 2019).

(14) Recursive structure in vowels and consonants: the CV unit [kʰi]



In (14) the ultimate head of the CV-sized structure is the highest |A|. It will be recalled from section 3.3 that a head |A| is pronounced as [ə] (i.e. the acoustic baseline) if no dependent elements are present. But in this case the head element |A| has a dependent |I|, and the acoustic signature of this dependent |I| overrides that of its head. As a result, the |A|-headed expression |A I| is realised as [i]. (For an explanation of how other vocalic expressions are realized, see Nasukawa 2016.)

Moving one structural level down, this |I| element now functions as a domain head and takes |H| as its dependent. The noise element |H| represents a range of obstruent-type properties including voicelessness and aspiration (see (2) above), indicating that everything below it in the structure refers to a consonant. As already noted, this consonantal domain (headed by |H|) cannot be realised simultaneously with the vocalic domain above it (headed by |A|) because the two domains involve articulatory gestures that are incompatible. At the next level of embedding the element |H| takes |U| as a dependent, where single |U| is realised as velar resonance (Nasukawa 2016, cf. Backley and Nasukawa 2009). Thus, the consonant is identified as a velar obstruent of some kind. This |U| then becomes a head, taking the edge element |ʔ| (defining occlusion) as its dependent. The consonant structure up to this point (i.e. $[[ʔ\ U]_U\ H]_H$) may be phonetically realised as a velar stop [k].

Finally, the lowest part of the structure contains a second token of |H|. As a dependent of |ʔ|, this |H| occupies the most deeply embedded part of the structure, which maximises its ability to carry linguistic information – recall that dependents are structurally unimportant but informationally rich. This means that this lower |H| is realised in its exaggerated (prominent) form, namely, as aspiration. Thus, the |H|-headed domain is interpreted as an aspirated velar plosive [k^h]. Together with the vocalic structure above it, the entire expression in (14) is realised as the CV-sized unit [k^hi]. This outcome is determined by the principle of phonetic interpretation in (15) (Nasukawa, Backley, Yasugi and Koizumi 2019).

(15) Type A (CV) precedence:
A domain located at a lower level (C domain) is phonetically realised before a domain located at a higher level (V domain).

This general principle, dubbed Type A, is observed in the vast majority of languages, including English. In (15) it is formulated in terms of a domain's position in the hierarchical element structure. But it may also be expressed by referring to the extent of the carrier signal's modulation: a domain associated with a bigger modulation (typically an obstruent consonant) precedes a domain with a smaller modulation (typically a vowel). This alternative way of interpreting (15) is based on the idea that domains located at the lower end of the hierarchy contain more linguistic information than those higher up, and it assumes that consonants tend to be richer in linguistic information than vowels (i.e. consonant representations employ a larger set of contrastive properties).

In the small number of languages which do not observe the Type A principle, we find a mechanism of phonetic realisation that is exactly the reverse of the one described in (15). The Mayan language Kaqchikel is one such system, in which the structure in (14) is predicted to have a VC realisation rather than CV. Following Nasukawa, Backley, Yasugi and Koizumi (2019), we assume that Kaqchikel adheres to the alternative principle of realisation in (16).

(16) Type B (VC) precedence:
A domain located at a higher level (V domain) is phonetically realised before a domain located at a lower level (C domain).

Expressed in terms of carrier signal modulations, Kaqchikel should display a pattern that is the opposite of the Type A pattern. That is, in the Type B pattern a domain associated with a bigger carrier signal modulation (typically an obstruent consonant) will be phonetically realised after a domain with a smaller modulation (typically a vowel).

The parametric difference between Type A precedence and Type B precedence (Nasukawa 2016) rests on the following two assumptions. First, all languages use the same hierarchical melodic structure, which is defined only by head-dependency relations between elements. And second, cross-linguistic variation is limited to whether a language uses a V-final or a V-initial precedence relation when phonological structure is phonetically realised. This typological variation is formalised as a parametric choice between (15) and (16). An explanation for why (15) is far more widespread than (16) may involve a discussion of physiological and psychological factors as well as purely linguistic factors. For the moment we leave this question open; further research will be needed to fully understand the general preference for the CV pattern over the VC pattern.

5 Summary

We have outlined a unified model of phonological structure which represents both melody and prosody by referring only to elements. In addition to performing their usual melodic functions, elements are projected upwards through the prosodic hierarchy to define successively wider prosodic domains. These domains replicate the standard prosodic units labelled 'nucleus', 'syllable', 'foot' and 'word'. The motivation for rejecting these traditional labels is that they are specific to just one part of structure, namely prosody. And this presents a problem, such that if melody and prosody are described using different sets of units, we fail to capture any non-arbitrary relation between them. We have also argued that a minimal 'syllable' contains just a bare head element, either |A|, |I| or |U|, selected by parameter. The choice of head element reflects the way a given language interprets empty nuclei.

References

Anderson, John and Colin J. Ewen. 1987. *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, John and Charles Jones. 1974. Three theses concerning phonological representations. *Journal of Linguistics* 10, 1–26.

Backley, Phillip. 2011. *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Backley, Phillip. 2012. Variation in Element Theory. *Linguistic Variation* 12(1), 57–102.

Backley, Phillip and Kuniya Nasukawa. 2009. Representing labials and velars: A single 'dark' element. *Phonological Studies* 12, 3–10.

Backley, Phillip and Kuniya Nasukawa. 2010. Consonant-vowel unity in Element Theory. *Phonological Studies* 13, 21–28.