Introduction: aims and content

Paul de Lacy

Introduction

Phonological theory deals with the mental representation and computation of human speech sounds. This book contains introductory chapters on research in this field, focusing on current theories and recent developments.

1 Aims

This book has slightly different aims for different audiences. It aims to provide concise summaries of current research in a broad range of areas for researchers in phonology, linguistics, and allied fields such as psychology, computer science, anthropology, and related areas of cognitive science. For students of phonology, it aims to be a bridge between textbooks and research articles.

Perhaps this book's most general aim is to fill a gap. I write this introduction ten years after Goldsmith's (1995) *Handbook of Phonological Theory* was published. Since then, phonological theory has changed significantly. For example, while Chomsky & Halle's (1968) *The Sound Pattern of English* (*SPE*) and its successors were the dominant research paradigms over a decade ago, the majority of current research articles employ Optimality Theory, proposed by Prince & Smolensky (2004). Many chapters in this book assume or discuss OT approaches to phonology.

Another striking change has been the move away from the formalist conception of grammar to a functionalist one: there have been more and more appeals to articulatory effort, perceptual distinctness, and economy of parsing as modes of explanation in phonology. These are just two of the many developments discussed in this book.

2 Website

Supplementary materials for this book can be found on the website: http://handbookofphonology.rutgers.edu.

3 Audience and role

The chapters are written with upper-level undergraduate students and above in mind. As part of a phonology course, they will serve as supplementary or further readings to textbooks. All the chapters assume some knowledge of the basics of the most popular current theories of phonology. Many of the chapters use Optimality Theory (Prince & Smolensky 2004), so appropriate background reading would be, for example, Kager's (1999) textbook *Optimality Theory*, and for the more advanced McCarthy's (2002) *A Thematic Guide to Optimality Theory*.

Because it is not a textbook, reading the book from beginning to end will probably not prove worthwhile. Certainly, there is no single common theme that is developed step-by-step throughout the chapters, and there is no chapter that is a prerequisite for understanding any other (even though the chapters cross-reference each other extensively). So, the best use of this book for the reader is as a way to expand his/her knowledge of phonology in particular areas after the groundwork provided by a textbook or phonology course has been laid.

This book is also not a history of phonology or of any particular topics. While it is of course immensely valuable to understand the theoretical precursors to current phonological theories, the focus here is limited to issues in recent research.

4 Structure and content

The chapters in this book are grouped into five parts: (I) conceptual issues, (II) prosody, (III) segmental phenomena, (IV) internal interfaces, and (V) external interfaces.

The 'conceptual issues' part discusses theoretical concepts which have enduring importance in phonological theory: i.e. functionalist vs. formalist approaches to language, markedness theory, derivation, representation, and contrast.

Part II focuses on the segment and above: specifically prosodic structure, sonority, and tone. Part III focuses on subsegmental structure: features and feature operations. The chapter topics were chosen so as to cover a wide range of phenomena and fit in with the aims of phonology courses. However, while the areas in Parts II and III are traditionally considered distinct, the boundaries are at least fluid. For example, Gussenhoven

(Ch.11) observes that research on tone and intonation seems to be converging on the same theoretical devices, so the tone–intonation divide should not be considered a theoretically significant division. In contrast, some traditionally unified phenomena may consist of theoretically distinct areas: Archangeli & Pulleyblank (Ch.15) observe that there may be two separate types of harmony that require distinct theoretical mechanisms. Nevertheless, the division into discrete phenomena is inevitable in a book of this kind as in practice this is how they are often taught in courses and conceived of in research.

Part IV deals with 'internal interfaces' – the interaction of the phonological component with other commonly recognized modules – i.e. phonetics (Kingston Ch.17), syntax (Truckenbrodt Ch.18), and morphology (Ussishkin Ch.19 and Urbanczyk Ch.20).

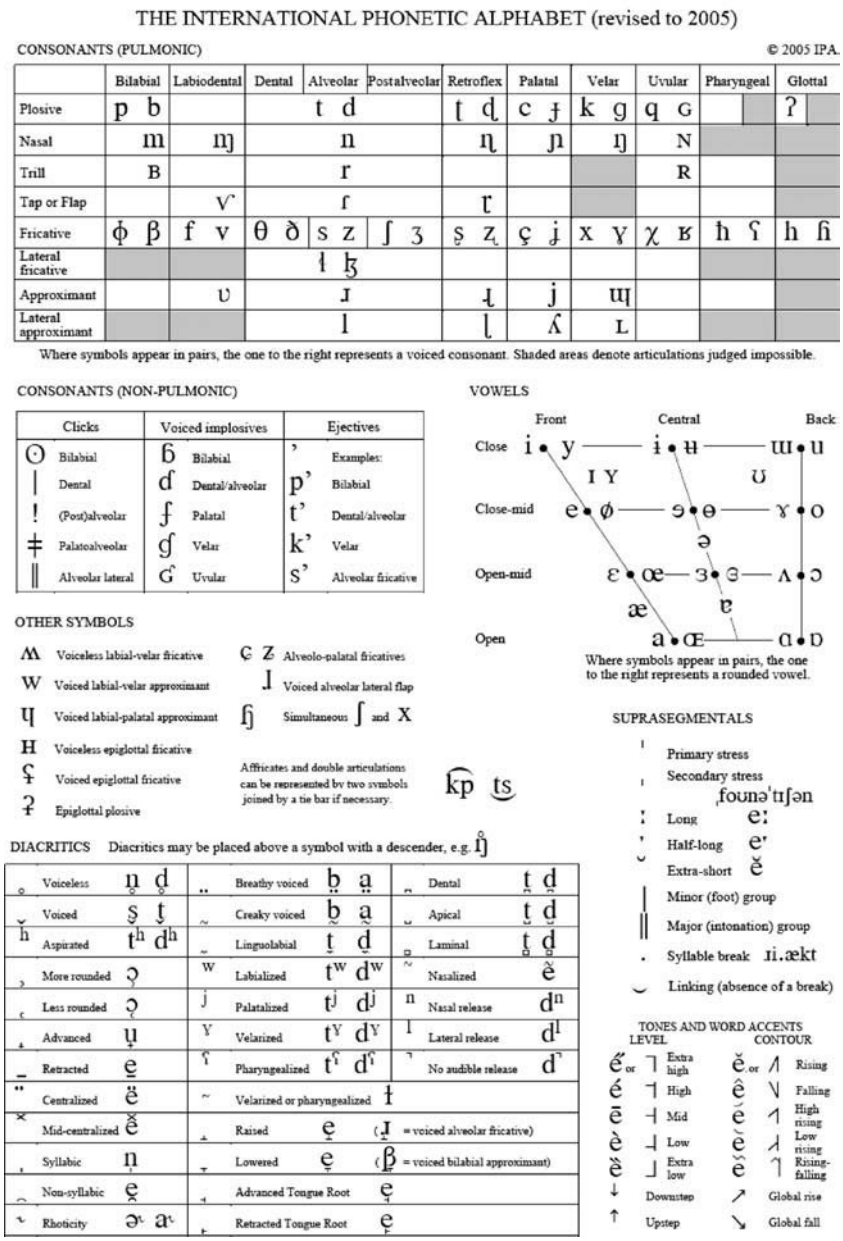
Part V focuses on a variety of areas that do not fit easily into Parts I–IV. These include well-established areas such as diachronic phonology (Bermúdez-Otero Ch.21), areas that have recently grown significantly (e.g. language acquisition – Fikkert Ch.23) or have recently provided significant insight into phonological theory (e.g. free variation – Anttila Ch.22, learnability – Tesar Ch.24, phonological impairments – Bernhardt & Stemberger Ch.25).

Practical reasons forced difficult decisions about what to exclude. Nevertheless, as a number of phonologists kindly offered their views on what should be included I hope that the topics covered here manage to reflect the current concerns of the field.

While phonological research currently employs many different transcription systems, in this book an effort has been made to standardize transcriptions to the International Phonetic Alphabet (the IPA) wherever possible:

http://www2.arts.gla.ac.uk/IPA/index.html.

Chart of the International Phonetic Alphabet (revised 1993, updated 1996)



This chart is provided courtesy of the International Phonetics Association, Department of Theoretical and Applied Linguistics, School of English, Aristotle University of Thessaloniki, Thessaloniki 54124, GREECE.

1

Themes in phonology

Paul de Lacy

1.1 Introduction

This chapter has two aims. One is to provide a brief outline of the structure of this book; this is the focus of Section 1.1.1. The other – outlined in Section 1.1.2 – is to identify several of the major themes that run throughout.

1.1.1 Structure

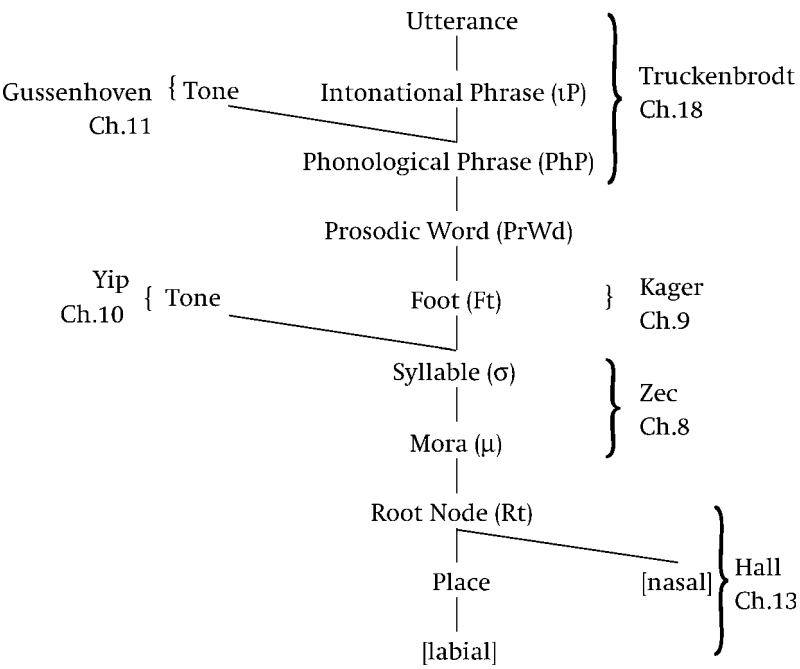
Several different factors have influenced the contents and structure of this Handbook. The topics addressed reflect theoretical concerns that have endured in phonology, but they were also chosen for pedagogical reasons (i.e. many advanced phonology courses cover many of the topics here). There were also 'traditional' reasons for some aspects of organization. While these concerns converge in the main, there are some points of disagreement. For example, there is a traditional distinction between the phonology of lexical tone and intonation, hence the separate chapters by Yip (Ch.10) and Gussenhoven (Ch.11). However, Gussenhoven (11.7) comments that theoretically such a division may be artificial.

Consequently, it is not possible to identify a single unifying theoretical theme that accounts for the structure of this book. Nevertheless, the topics were not chosen at random; they reflect many of the current concerns of the field. In a broad sense, these concerns can be considered in terms of representation, derivation, and the trade-off between the two. 'Representation' refers to the formal structure of the objects that the phonological component manipulates. 'Derivation' refers to the relations between those objects.

Concern with representation can be seen throughout the following chapters. Chomsky & Halle (1968) (*SPE*) conceived of phonological representation as a string of segments, which are unordered bundles of features. Since then, representation has become more elaborate. Below the segment, it is widely accepted that features are hierarchically organized (see discussion

and references in Hall Ch.13). Above the segment, several layers of constituents are now commonly recognized, called the 'prosodic hierarchy' (Selkirk 1984b). Figure (1) gives a portion of an output form's representation; it categorizes the chapters of this book in terms of their representational concerns. There is a great deal of controversy over almost every aspect of the representation given below – Figure (1) should be considered a rough expositional device here, not a theoretical assertion; the chapters cited should be consulted for details.

(1)



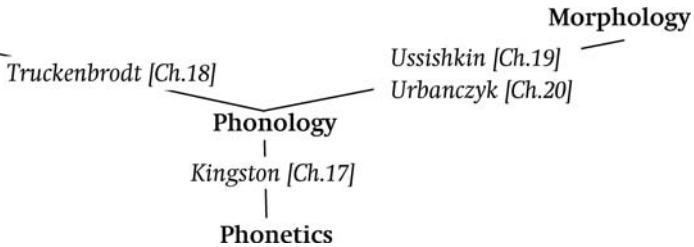
Harris (Ch.6) should be added to the chapters cited in (1); Harris' chapter is concerned with broader principles behind representation, including the notion of constituency, whether certain sub-constituents are phonologically prominent (i.e. headedness), and hierarchical relations.

Not represented in (1) is the interaction between constituents. For example, de Lacy (Ch.12) examines the interaction of tone, the foot, and segmental properties. Similarly, a part of Kager (Ch.9) is about the relation between the foot and its subconstituents. At the segmental level, three chapters are concerned with the interaction of segments and parts of segments: Baković (Ch.14), Archangeli & Pulleyblank (Ch.15), and Alderete & Frisch (Ch.16). For example, Baković's chapter discusses the pressure for segments to have identical values for some feature (particularly Place of Articulation).

Figure (2) identifies the chapters that are concerned with discussing the interaction of different representations. For example, Truckenbrodt (Ch.18) discusses the relation of syntactic phrases to phonological phrases. Ussishkin (Ch.19) and Urbanczyk (Ch.20) do the same for the relation of morphological

and phonological structure. Kingston (Ch.17) discusses the relation of phonological to phonetic structures.

(2)



There is also a 'derivational' theme that runs through the book chapters. McCarthy (Ch.5) focuses on evidence that there are relations between morphologically derived forms, and theories about the nature of those relations. Discussion of derivation has traditionally focused on the relation between input and output forms, and between members of morphological paradigms. However, the traditional conception of derivation has been challenged in Optimality Theory by McCarthy & Prince's (1995a, 1999) Correspondence Theory – the same relations that hold between separate derivational forms (i.e. input∼output, paradigmatic base∼derivative) also hold in the same output form between reduplicants and their bases; thus Urbanczyk's (Ch.20) discussion of reduplication can be seen as primarily about derivation, in this broadened sense.

Of course, no chapter is entirely about the representation of constituents; all discuss derivation of those constituents. In serialist terms, 'derivation of constituents' means the rules by which those constituents are constructed. In parallelist (e.g. Optimality Theoretic) terms, it in effect refers to the constraints and mechanisms that evaluate competing representations.

There is a set of chapters whose primary concerns relate to both representation and derivation: Prince (Ch.2), Gordon (Ch.3), Rice (Ch.4), and Steriade (Ch.7) discuss topics that are in effect meta-theories of representation and derivation. Gordon (Ch.3) examines functionalism – a name for a set of theories that directly relate to or derive phonological representations (and potentially derivations) from phonetic concerns. Rice (Ch.4) discusses markedness, which is effectively a theory of possible phonological representations and derivations. Steriade (Ch.7) discusses the idea of phonological contrast, and how it influences representation and derivation.

Rice's discussion of markedness makes the current tension between representation- and derivation-based explanations particularly clear. Broadly speaking, there have been two approaches to generalizations like "an epenthetic consonant is often [ʔ]". One assigns [ʔ] a representation that is different (often less elaborate) than other segments; the favouring of epenthetic [ʔ] over other segments is then argued to follow from general derivational principles of structural simplification. The other is to appeal to derivational principles such as (a) constraints that favour [ʔ] over every other segment and (b) no constraint that favours those other segments over

[ʔ]; [ʔ] need not be representationally simple (or otherwise remarkable) in this approach. These two approaches illustrate how the source of explanation – i.e. derivation and representation – is still disputed. The same issue is currently true of subsegmental structure – elaborated derivational mechanisms may allow simpler representational structures (Yip 2004).

Part V of this book contains a diverse array of phonological phenomena which do not fit easily into the themes of representational and derivational concerns. Instead, their unifying theme is that they are all areas which have been the focus of a great deal of recent attention and have provided significant insight into phonological issues; this point is made explicitly by Fikkert (Ch.23) for language acquisition, but also applies to the other areas: diachronic phonology (Bermúdez-Otero Ch.21), free variation (Anttila Ch.22), learnability (Tesar Ch.24), and phonological disorders (Bernhardt & Stemberger Ch.25). There are many points of interconnection between these chapters and the others, such as the evidence that phonological disorders and language acquisition provide for markedness.

Standing quite apart from all of these chapters is Prince (Ch.2). Prince's chapter discusses the methodology of theory exploration and evaluation.

In summary, no single theoretical issue accounts for the choice of topics and their organization in this book. However, many themes run throughout the chapters; the rest of this chapter identifies some of the more prominent ones.

1.1.2 Summary of themes

One of the clearest themes seen in this book is the influence of Optimality Theory (OT), proposed by Prince & Smolensky (2004).[1] The majority of chapters discuss OT, reflecting the fact that the majority of recent research publications employ this theory and a good portion of the remainder critique or otherwise discuss it.[2] However, one of the sub-themes found in the chapters is that there are many different conceptions and sub-theories of OT, although certain core principles are commonly maintained. For example, some theories employ just two levels (the input and output), while others employ more (e.g. Stratal OT – McCarthy 5.4). Some employ a strict and totally ordered constraint ranking, while others allow constraints to be unranked or overlap (see Anttila 22.3.3 and Tesar 24.4 for discussion). Theories of constraints differ significantly among authors, as do conceptions of representation (see esp. Harris Ch.6).

Another theme that links many of the chapters is the significance of representation and how it contributes to explanation. The late 1970s and 1980s moved towards limiting the form of phonological rules and elaborating the representation by devices such as autosegmental association, planar segregation, lack of specification, and feature privativity. In contrast, Harris (6.1) observes that the last decade has seen increased reliance on constraint form and interaction as sources of explanation. Constraint

interaction as an explanatory device appears in many of the chapters. Section 1.3 summarizes the main points.

Section 1.4 discusses the increasing influence of Functionalism in phonology, a theme that is examined in detail by Gordon (Ch.3). Reference to articulatory, perceptual, and parsing considerations as a source of phonological explanation is a major change from the Formalist orientation of *SPE* and its successors. This issue recurs in a number of chapters, some explicitly (e.g Harris 6.2.2, Steriade 7.5), and in others as an implicit basis for evaluating the adequacy of constraints.

Of course, the following chapters identify many other significant themes in current phonological theory; this chapter focuses solely on the ones given above because they recur in the majority of chapters and are presented as some of the field's central concerns.

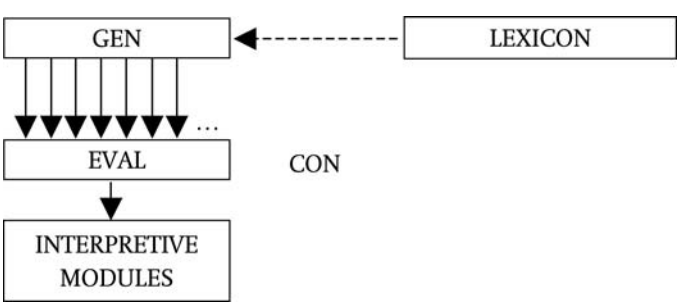
1.2 The influence of Optimality Theory

Optimality Theory is explicitly discussed or assumed in many chapters in this volume, just as it is in a great deal of current phonological research ('current' here refers to the time of writing – the middle of 2005). This section starts by reviewing OT's architecture and core properties. The following sections identify particular aspects that prove significant in the following chapters, such as the notion of faithfulness and its role in derivation in Section 1.2.1, some basic results of constraint interaction in Section 1.2.2, and its influence on conceptions of the lexicon in Section 1.2.3. The sections identify some of the challenges facing OT as well as its successes and areas which still excite controversy. The relation of OT to other theories is discussed in Section 1.2.4.

OT Architecture

OT is a model of grammar – i.e. both syntax and phonology (and morphology, if it is considered a separate component); the following discussion will focus exclusively on the phonological aspect and refer to the model in (3).

(3) *OT architecture*



For phonology, the GEN(erator) module takes its input either directly from the lexicon or from the output of a separate syntax module. GEN creates a possibly infinite set of candidate output forms; the ability to elaborate on the input without arbitrary restraint is called 'freedom of analysis'. In Prince & Smolensky's original formulation, every output candidate literally contained the input; to account for deletion, pieces of the input could remain unparsed (i.e. not incorporated into prosodic structure) which meant they would not be phonetically interpreted. Since McCarthy & Prince (1995a/1999), the dominant view is that output candidates do not contain the input, but are related to it by a formal relation called 'correspondence'; see Section 1.2.1 for details (cf. Goldrick 2000).

One significant restriction on GEN is that it cannot alter the morphological affiliation of segments ('consistency of exponence' – McCarthy & Prince 1993b). In practice it is common to also assume that GEN requires every output segment to be fully specified for subsegmental features, bans floating (or 'unparsed') features (except for tone – Yip 10.2.2, Gussenhoven 11.5.1), and imposes restrictions on the form of prosodic and subsegmental structure (though in some work they are considered violable – e.g. Selkirk 1995a, Crowhurst 1996, cf. Hyde 2002).

The EVAL(uator) module determines the 'winner' by referring to the constraints listed in CON (the universal constraint repository) and their language-specific ranking. Constraints are universal; the only variation across languages is (a) the constraints' ranking, and (b) the content of the lexicon. The winner is sent to the relevant interpretive component (the 'phonetic component' for phonology – Kingston Ch.17).

There are two general types of constraint: Markedness and Faithfulness. Markedness constraints evaluate the structure of the output form, while Faithfulness constraints evaluate its relationship to other forms (canonically, the input – see McCarthy Ch.5).[3] As an example, the Markedness constraint ONSET is violated once for every syllable in a candidate that lacks an onset (i.e. every syllable that does not start with a non-nuclear consonant – Zec 8.3.2). [ap.ki] violates ONSET once, while [a.i.o] violates it three times. The Faithfulness constraint I(nput)O(utput)-MAX is violated once for every input segment that does not have an output correspondent: e.g. /apki/ → [pi] violates IO-MAX twice (see Section 1.2.1 for details).

In each grammar the constraints were originally assumed to be totally ranked (although evidence for their exact ranking may not be obtainable in particular languages); for alternatives see Anttila (Ch.22). Constraints are violable; the winner may – and almost certainly will – violate constraints. However, the winner violates the constraints 'minimally' in the sense that for each losing candidate L, (a) there is some constraint K that favors the winner over L and (b) K outranks all constraints that favor L over the winner (a constraint 'favors' *x* over *y* if *x* incurs fewer violations of it than *y*); see Prince (2.1.1) for details.