

- Ohala, J. J. (1981), 'Speech Timing as a Tool in Phonology', *Phonetica*, 38: 204–12.  
 — (1983), 'The Origin of Sound Patterns in Vocal Tract Constraints', in P. F. MacNeilage (ed.), *The Production of Speech* (New York: Springer), 189–216  
 — (1990), 'There is no Interface between Phonology and Phonetics: A Personal View', *Journal of Phonetics*, 18/2: 153–71.  
 — (1995), 'The Perceptual Basis of Some Sound Patterns', in B. Connell and A. Arvaniti (eds.), *Papers in Laboratory Phonology IV* (Cambridge: Cambridge University Press), 87–92.  
 Peirce, C. (1933), *Collected Papers* (Cambridge, Mass.: Harvard University Press).

## 2

---

## The Ontology of Phonology (Revised)

SYLVAIN BROMBERGER AND MORRIS HALLE

### 1. PHONOLOGY AND THE PHILOSOPHY OF LANGUAGE

Though many philosophers of language have views on empirical linguistics, few, if any, have given serious attention to phonology. Recent anthologies and books on the philosophy of language either do not mention phonology at all, or at best perfunctorily restate crude and outdated notions on the subject. This is somewhat surprising, since the facts that phonology studies are critical to the individuation of expressions and to their character as objects of speech perception or outputs of speech production. But for these facts, there would be no syntax or semantics of natural languages besides sign languages (also neglected by philosophers), and philosophers deliberating about such languages would have to be silent.

There are a number of explanations for this neglect. To begin with, recent philosophers of language generally belong to an intellectual tradition that admits no essential differences between natural languages and some of their contrived extensions. This was pointed out a long time ago by Strawson (1950), though he had other shortcomings in mind. Philosophic discussions thus generally abstract not only from differences between English, German, Japanese, and other natural languages, but also from differences between these real languages and notational systems used in mathematics, logic, physics, chemistry, biology, linguistics, etc. Such notational systems do have a syntax (albeit usually one that has very little in common with the syntax of natural languages), a semantics, and a pragmatics, but happen to have no phonology. Their minimal units are normally ideographs that encode wordlike units—rather than phonetic or even orthographic ones—open to many phonologically unrelated pronunciations (if pronounceable at all). Nothing in such notational systems corresponds to the phonologies of natural languages, and nothing about them can thus be captured in an overarching phonological

Chapter 2 is based on chapter 9 of *On What We Know We Don't Know*, by Sylvain Bromberger, copyright © 1992 by CSLI Publications, Stanford University, Stanford, CA 94305. The Publisher gratefully acknowledges the permission of CSLI Publications to include the revised version in this book.

We are grateful to Ned Block, Nancy Bromberger, George Boolos, Noam Chomsky, Leonard Clapp, Alec Marantz, Wayne O'Neil, Jean Michel Roy, and Linda Wetzel for comments on previous drafts of this chapter.

doctrine linked to the overarching semantic and syntactic doctrines studied by philosophers. So it is not surprising that, though Frege and his successors include signs in their Sign–Sense–Nominatum triad, they have nothing of interest to say about signs as things uttered and heard. Even philosophers who focus primarily on natural languages belong to that tradition and do not discriminate between aspects peculiar to real languages—that is, articulated languages whose inherently spoken tokens necessarily vanish as soon as produced, and aspects peculiar to conventional notational systems (including writing systems) with characteristically enduring tokens.

Furthermore, philosophers generally seem to believe that there cannot be anything of philosophic interest about phonology. This attitude flows naturally from the previous one. The ideographs used by scientists are adopted through open and explicitly published conventions that determine everything true of them qua signs. Since they are also semantically wordlike, or at least morpheme-like, their subsegments have no autonomous status and raise no philosophic problems. How could there be anything of philosophic interest about the shape of simple numerals, or about the horizontal segment in an inverted V in a quantifier, or in the vertical line of the F for force? It is easy to pass from this outlook to the view that there cannot be anything of philosophic import about the spelling of words, and thence to the view that there cannot be anything of philosophic import about their pronunciation. And is not phonology ‘just’ about pronunciation?

Whatever the explanation for philosophers’ neglect of phonology, we think that it has a cost. To begin with, we think that it is a mistake to lump all lexical systems together as forming some kind of natural family open to ready generalizations. It glosses over too many crucial differences, and attention to phonology can highlight important ones. Furthermore, we think that no theory about the relation between natural language signs and their referents (or their meaning, or their use) can be trustworthy that nonchalantly takes the nature of those signs for granted. Finally, we think an adequate understanding of the ontology of language—of the objects whose existence constitutes the reality of language—must include an adequate conception of the objects investigated by phonology. More specifically, an adequate conception of language must check our tendency, when we reflect about language, to slip thoughtlessly between talk about individual utterances and talk about types, as if such slips were always innocuous and easily fixed ways of avoiding pedantry. Spoken or subvocally produced tokens are transitory events that are finite in number, that occur in time and space, that can be perceived, that are shaped by their speaker’s occurrent intentions, and that are subject to norms fixed in their speakers’ mental make-up. Types—if there are types—are abstract entities, neither in time nor in space, devoid of causal histories or causal consequences, hence beyond perception. Types—if there are types—outdistance tokens, in number (since there are infinitely many of them). Tokens and types—if there are types—are thus utterly different. Conflating them, or taking their connection as given, is bound to produce confusions and incoherence. But giving

each its due, and understanding their connection, will not be possible unless we see how the type/token distinction fares in phonology.

## 2. THE ISSUES

Linguistics, like Gaul, is traditionally divided into three parts, syntax, semantics, and phonology, the latter being presumably concerned with the sound aspect of language.<sup>1</sup> The issues we plan to discuss in this chapter concern phonology directly and the other two branches indirectly. We take phonological theory to be about the world, about reality, and thus about certain items in the world—certain ‘particulars’, as metaphysicians<sup>2</sup> might put it—whose existence is attested to by the fact that people speak. What is the nature of these ‘particulars’? In the first part of our chapter, we will address that question. Our answer will be that phonology is about concrete mental events and states that occur in real time, in real space, have causes, have effects, are finite in number, in other words are what metaphysicians would call CONCRETE PARTICULARS closely linked to, but distinct from, those described by traditional phoneticians concerned with articulatory or acoustic events. In the second part of our chapter we will consider a very different answer, according to which phonology is about types, a certain species of abstract, causally impotent, non-spatiotemporal entities, possibly infinite in number, and distinct from real live utterances.<sup>3</sup> Phonology, like the rest of linguistics, is normally expounded as if it were about types. But does this mean that the discipline is committed to there being such abstract entities as types? We will argue that it is not.

<sup>1</sup> This chapter is a somewhat modified version of a paper originally given in Uppsala, Sweden, at the Ninth International Congress of Logic, Methodology, and Philosophy (Bromberger and Halle 1994). All the references to actual utterances are to utterances produced by Sylvain Bromberger on that occasion. Readers should keep in mind that the chapter was intended for an audience of philosophers of language and philosophers of science—that is, an audience largely unfamiliar with phonology—and that one of our aims was to get such philosophers to appreciate the conceptual and ontological issues raised by phonology. The present version is still mostly untechnical and suited for an audience that knows little about phonology but is interested in the nature of language and its study. Since many phonologists, in practice if not in theory, write and talk as if the objects of their investigations were abstract entities—the sort of entities normally called types—the issues raised here will be pertinent to their interests as well. For a paper on how these issues arise in the debate between rule-based phonology and optimality theory, see Bromberger and Halle (1997).

<sup>2</sup> See Bromberger and Halle (1989) on why phonology is fundamentally different from syntax and semantics. We take morphology to be at the intersection between syntax and phonology.

<sup>3</sup> The type/token distinction, in the sense relevant to our discussion, is usually traced to a puzzling passage by C. S. Peirce (1933: iv, 423): ‘... there is but one word “the” in the English language; and it is impossible that this word should lie visibly on a page or be heard in any voice, for the reason that it is not a Single thing or Single event. It does not exist: it only determines things that do exist. Such a definitely significant Form, I propose to term a TYPE. A single event which happens once and whose identity is limited to that happening or a Single object or thing which is in some single place at any one instant of time, such event or thing being significant only as occurring just when and where it does, such as this or that word on a single line of a single page of a single copy of a book, I will venture to call

In the course of our discussion we will use some technical notation but we will keep it to a minimum and will explain it as we go. We will also talk from within a framework that some linguists may reject. That cannot be helped. Linguistics is in constant flux and full of controversies. Nothing of real interest in it is conclusively established once and for all and to everybody's satisfaction.

### 3. ON THE NATURE OF TOKENS

#### 3.1. The phonological representation of a token

Let us begin by thinking about spoken tokens. And to fix ideas, let us focus on a very specific one, the one that I, the speaker,<sup>4</sup> will now produce:

(1) *The merchant sold shelves.*

That token is now history! Only time travel could enable us ever to hear it again. It could, of course, be duplicated, but it itself is gone forever. We will come back to the fact that it could be duplicated, but let us forget about that right now. Let us concentrate on the specific event that happened a few seconds ago. We will refer to it as event (1) since we are not able to display it again.

Actually many things happened when event (1) occurred. That is why it could be studied by more than one discipline and be analysed differently by each. So, for instance, noises happened, and event (1) therefore could be investigated under acoustics and given an acoustical analysis. Bodily movements happened, and event (1) could therefore be studied under motor behaviour and given an articulatory analysis. Brain and neurological events happened, and (1) could be looked at under neurology and given a neurological analysis. And so on.

However, we exhibited event (1) to illustrate a phonological event that is an event that can be examined in the light of phonological theory and given a phonological analysis.

What would such an analysis tell us about (1)?

Well, let us look at how phonologists would represent (1) in the notation of phonology.

They would represent it as follows (the dots between (2a), (2b), and (2c) represent 'lines' that we omit for present purposes):

TOKEN. These tokens are actual utterances produced by specific persons, at specific times, at specific places. Types, if there are types, are not. It is important in this connection to distinguish tokens from occurrences of types. Thus the sentence type *Joe's cat hates Mary's cat* contains two occurrences of the type *cat* but contains no tokens, since types are not made up of tokens. For other discussions and attempts to clarify the distinction see Cartwright (1987), Quine (1987), Bromberger (1989), Hutton (1990), Katz (1990), Wetzel (1993), Burton-Roberts (1994, and this volume), Szabo (1999).

<sup>4</sup> We use the first-person singular to mention Sylvain Bromberger as producer of tokens displayed during the talk in Uppsala and use the first-person plural to mention ourselves, the two authors. We use script font to indicate displayed actual specific spoken tokens. Readers of the chapter should thus keep in mind that our use of script font is to point to episodes that they can no doubt imagine but that took place in Uppsala when this chapter was orally presented.

(2) (a) {[ðə], Art ... } + {[mɜrtʃənt], Noun ... } + {Q, Sing ... } + {[sɛl], Vb ... } + {Q, Past ... } + {[ʃɛlv], Noun ... } + {Q, Plur ... }

(b) {[ðə], Art ... } + {[mɜrtʃənt], Noun ... } + {Q, Sing ... } + {[sɔl], Vb ... } + {Q, Past ... } + {[ʃɛlv], Noun ... } + {Q, Plur ... }

(c) ðəmɜ:ʃntʃoldʃɛlv z

In other words, they would represent it as a sequence of lines, a DERIVATION. Each line would purport to stand for some fact about (1), and the ordering would purport to stand for further facts about it.

What kind of facts?

We are going to go through the derivation step by step to answer that question. But before doing so, we want to describe the general character of that answer.

My production of event (1) was an action. Like other actions, it was therefore brought about by a distinctive kind of mental set—something we will call an INTENTION. But this term, as we use it, is not to be taken altogether literally. We use it to refer to a familiar kind of purposive mental stance. Think of someone aiming a rifle at a target. That person moved and positioned limbs, head, eyes, and so on in certain ways. But more went on. After all, the movements were not made accidentally, or by way of checking whether the barrel is in line with the butt. The person was set psychologically in a distinct way—that is, had distinct intentions. More specifically, a person who aims a rifle has certain effects in mind, plans moves in ways calculated to achieve those effects, and, crucially from our point of view, has the intellectual capacity to select those effects and to devise the gestures that achieve them. The uttering of (1), like the aiming of a rifle, also required a distinctive mindset, distinctive intentions on my part, intentions that I could not have formed without certain pre-existing intellectual capacities. Of course, I had many intentions when I produced it: I intended to give you an example, I intended to be understood, I intended to produce a sentence that you have probably never heard before. But only some of my intentions account for the fact that I PRONOUNCED (1), that (1) was an action of pronouncing something in a language I know, in my idiolect of English. Those intentions are the kinds of facts about (1) that we take (2) to represent.

#### 3.2. The last line of the derivation

Let us now look at the last line of the derivation (2), that is (2c).

(2c) could be construed as a phonetic transcription of the utterance (1). Formally it is a string of letters from an alphabet in which each letter traditionally stands for a speech sound.<sup>5</sup> Speech sounds, as articulatory events, are not

<sup>5</sup> The fact that the utterance consists of sequences of discrete sounds is the insight on which all alphabetic writing systems are based. It may, therefore, appear to be self-evident. Yet, when we last asked our colleagues working on the automatic analysis of speech, we were told that no one has yet found a reliable mechanical procedure that can segment any arbitrary utterance into its constituent sounds.

unanalysable entities. They are rather complexes of (phonetic) FEATURES. Thus each letter in (2c) stands for a particular complex of features. In (3) we have given a partial list of the feature composition of some of the component sounds of English.

(3)	p	b	m	f	v	t	d	n	s	z	k	g	
	-	-	-	+	+	-	-	-	+	+	-	-	continuant
	-	-	+	-	-	-	-	+	-	-	-	-	nasal
	-	+	+	-	+	-	-	+	-	+	-	+	voiced
			labial				coronal				dorsal		major articulator

You will readily notice that the three sets of consonants in (3) differ from each other in that each involves action by a different major articulator; that is, [pbmfv] are produced with active involvement of the lips; [tdnsz], with that of the tongue blade or coronal articulators; and [kg], with that of the tongue body or dorsal articulator. It is the major articulator that stops the airflow from out of the mouth in [-continuant] sounds, but allows it in the sounds that are [+continuant].

In addition to the major articulators the production of consonants involves other articulators as well. In particular, the consonants *m* and *n* are produced with a lowering of the velum, which allows air to flow through the speaker's nasal cavities exciting thereby the resonances of these cavities. In all other consonants, the velum is raised, no air flows through the nasal passages, and their characteristic resonances are not excited. This information is reflected by the pluses and minuses in the second line of (3). The third line reflects the behaviour of the vocal cords. It is the vocal cords that implement the feature [voiced]. They vibrate in [+voiced] sounds such as [bmvdnzɡ] and they are stationary in the [-voiced] consonants [pftsk].

We said a moment ago that (2c) *could* be construed as a phonetic transcription—that is, as a record of articulator movements and positionings. However, that is not the way we construe it! We construe (2c) as standing for a series of intentions that generated those movements. Each letter in (2c) stands for such an intention, and each of these intentions called for an arrangement of articulators in the expectation of distinctive auditory effects. Each was also the intention to act so as to produce a specific English SPEECH sound, and thus required a capacity that I acquired when I acquired English.

Consider, for instance, the *m* in (2c). It represents an intention (at the time) that called for simultaneously closing my mouth at the lips, lowering my velum, adjusting the stiffness of my vocal folds, and thereby producing a sound *m*. That is what the expansion in feature notation reflects. However, it does not represent an intention that merely called for going through all that gymnastics to produce the sound *m*. I could have intended that much without intending to produce an English sound—for instance while intending to hum a tune, or to imitate a cow, or to express mild surprise. The *m* in (2c) represents an intention to act so as to

produce a specific ENGLISH SPEECH SOUND. And I could not have formed that intention, that mindset, had I not acquired English.<sup>6</sup>

The other letters in (2c) stand for similar intentions to utter speech sounds in (1). Let us call them PHONETIC INTENTIONS.<sup>7</sup> What (2c) represents is, therefore, totally unlike what an oscillograph hooked to a microphone might have recorded, and this not only because some information recoverable from oscillograph records such as loudness, rate of speaking, fundamental frequency, and other characteristics of the speaker's voice cannot be inferred from (2c), but crucially because it stands for a different kind of event altogether. (2c) stands for the occurrence of phonetic intentions. Oscillographs hooked to microphones record the occurrence of noises.

But why not construe (2c) as standing for the actual movements that produced the noises rather than for mere intentions? The symbols, after all, were introduced in the discipline for that purpose! We have at least two reasons. The first is conceptual. (2c), as we shall see in a moment, represents the result of a mental computation. And we do not think that movements are necessary results of the kind of computation involved here. Results of such computations have content. (2c) happens to characterize a content that was executed, but need not have been executed.<sup>8</sup> It would not have been executed had we been reading or mulling subvocally. The second is empirical. When we execute a speech action we correct for all sorts of momentary impediments and conditions. (2c) says nothing about such corrections. It contains only linguistic information and takes into account only linguistic knowledge. As a description of the articulator actions, it might be false. So we use the traditional symbols, but we do not subscribe to their standard phonetic interpretation.<sup>9</sup>

### 3.3. The first line of the derivation

Let us now turn to (2a), the first line in the derivation. It represents another series of intentions responsible for event (1)—namely, the intentions to use certain words, for instance, the noun *merchant*, the verb *sell* marked for past tense, and so

<sup>6</sup> Compare with the availability of click sounds as phonemes to speakers of Bantu, but only as noises to speakers of English.

<sup>7</sup> For a related position see Liberman and Mattingly (1985, 1989), and Bromberger and Halle (1986), and Halle (1997).

<sup>8</sup> Elsewhere (Bromberger and Halle 1986a) we have called this kind of intentional content a *score*, to mark the analogy with a musical score, something that can be executed through motions, but need not be executed, and often is not. Inner discourse probably stops at the formation of such scores.

<sup>9</sup> Our interpretation of event (1) commits us to the occurrence of the mental events we call intentions over and above acoustical events and articulatory events. We present some of our reasons in Bromberger and Halle (1989). Lenneberg (1967) pointed out a long time ago that the neural paths to the various articulators implicated in the production of a speech sound are of different lengths and that instructions to move them must therefore leave the brain at different times, thus providing possible evidence for the existence of brain events corresponding to such intentions.

on, in a certain order. This is reflected, for instance, in (2a) by the clustering of phonetic symbols into larger bracketed wordlike units.

Forming the intention to produce (1) clearly required that I know the words I used, and that I retrieve them from memory. So, before discussing in more detail how (2a) relates to (1), let us look at how some linguists represent knowledge of words.

None of us is born with the knowledge of the words of our native language. That children learn words as they develop is obvious and moreover massively documented, and we all know that the process goes on through life: most of us have only recently acquired such words as *intifada*, *glasnost*, *scud*. The proposition that an essential part of acquiring a language consists in storing in one's memory (something representable as) a list of words, something we will call the Vocabulary, is therefore one of the most securely founded in all linguistics.

Many words that any speaker of English knows are complex in the sense that they incorporate affixes of various kinds. We illustrate this in (4).

- (4) (a) shelv-es, child-ren, bough-t, sol-d  
 (b) pre-dis-pose, un-happy, in-secure  
 (c) un-poison-ous-ness, ex-pre-sid-ent, contra-in-dic-ate-d  
 (d) kibbutz-im, hassid-im

In linguistics the term STEM is used to designate the element to which an affix is added and the term MORPHEME is used as a cover term for both affixes and stems. Like stems, affixes too must be learned and committed to memory (see (4d)).

A speaker's knowledge of morphemes can thus be represented as a list of items containing information about each morpheme stored in memory.

What information?

Obviously information about its meaning, its functional structure, and the thematic roles it assigns. Also about its lexical category—that is, whether it is a noun, verb, preposition, adjective, conjunction. And certainly information pertaining to how the morpheme is pronounced—that is, phonological information. All this can be thought of as encoded in a COMPLEX SYMBOL, made up of elements that stand for meaning, lexical category, and so on. The markers pertaining to how the morpheme is pronounced are of particular interest to us here. We will refer to each of them as the IDENTIFYING INDEX. The Vocabulary as a whole can thus be represented as a long list of such complex symbols, each of which contains, among other things, an identifying index.<sup>10</sup>

We now turn to the information in identifying indices.

<sup>10</sup> In talk about identifying indices we must take care to distinguish between what we presume to be 'in the mind' of the knower and representations in the notation of our theory. We use the term IDENTIFYING INDEX primarily to refer to representations in the notation of the theory, not what is 'in the mind'. We do, however, expect that research will eventually reveal elements of the mind/brain corresponding to these representations, and we may therefore sometimes avoid circumlocution by using the term and others like it to refer to these putative elements.

Most morphemes take on the same phonetic form regardless of syntactic and/or morphological contexts. Thus the verb *hint* shows up in the phonetic form representable as [hint] whenever uttered. So that string of phonetic symbols is used as its identifying index.

Other morphemes assume different phonetic forms depending on syntactic and/or morphological contexts. For instance, the stems *sell* and *shelf* were pronounced differently in (1) than in the following utterance, which I now produce:

(5) *The merchant sells a shelf.*

The identifying index of such stems is also a string of phonetic symbols—namely, [sel] and [ʃelf] in the two cases at hand. We will come back to why those particular strings.

Not only do some morphemes, notably the English plural and past-tense affixes, assume different phonetic forms in different contexts, but these forms can also be utterly dissimilar, and sometimes they do not appear phonetically at all!

So note what happens to the plural affix in the following cases:

- (6) cat/s, child/ren, kibbutz/im, alumna/i, stigma/ta, geese, moose

and to the past-tense morpheme in the following:

- (7) bake/d, playe/d, dream/t, sol/d, sang, hit

Halle (1990) has dubbed morphemes such as the Plural and Past morphemes, which behave in this very irregular fashion, ABSTRACT MORPHEMES and he has used Q, a symbol that has no direct phonetic interpretation, as their identifying index.<sup>11</sup>

With all this in mind, let us look again at (2a).

(2a) is a sequence of complex symbols each made up of an identifying index and other grammatical markers, all copied from the Vocabulary.

What facts about event (1) does (2a) represent?

(2a), as we said before, represents the intention to use certain words, but we can now be more explicit. (2a) represents the fact that (1), besides being produced by my phonetic intentions in (2c) was also produced by my intention to use the words retrieved from my memory whose identifying indices (and lexical category) appear in (2a).

But what are the phonetic symbols doing in (2a)? Take the initial m in the identifying index of *merchant*. Does it represent an intention, already present at that stage so to say, to produce a token of the phoneme m—that is, to close my lips,

<sup>11</sup> Other elements in the lexicon may also lack specific phonological content—for instance, PRO, empty complementizers, case, expletives, etc.—yet not be represented by a complex symbol that includes any identifying index like Q. Whether or not some morpheme must be represented with an identifying index like Q or no identifying index at all is a contingent matter to be settled on empirical grounds. Halle's proposal is not that we adopt a convention for the sake of giving all complex symbols a common format. It embodies the claim that all unarticulated morphemes are not phonologically equal.

lower my velum, slacken my vocal folds, and so on? Offhand that may seem reasonable. But consider then the  $\epsilon$  in the identifying index of *sell*. It cannot stand for an intention to produce a token of the phoneme  $\epsilon$ , to undertake the relevant gestures. No such intention was executed in producing (1). Could I have changed my mind between the times I picked the words and pronounced them? That strikes us as a cute but vacuous idea, too literal minded about our use of INTENTION.<sup>12</sup> As we see it, the role of the phonetic symbols in (2a) and in (2c) is very different. In (2a) they play a computational role. Formulae such as (2a) have two functions. On the one hand, they model an event, represent aspects of that event. On the other hand they are used to compute other formulae IN THE FORMALISM OF OUR THEORY. Phonetic symbols appear in (2a) essentially to simplify computations within the theory. In (2c) they have that role, but they also represent phonetic intentions. These roles, though connected, are different.<sup>13</sup>

Note that in the Vocabulary phonetic symbols could not stand for intentions either. The Vocabulary is a representation not of intentions, but of knowledge. But its formulae too enter into computations.

#### 3.4. The second line of the derivation

Let us now look at (2b).

(2b) stands between (2a) and (2c). It is like (2a), except that some of the phonetic symbols in the identifying indices have been changed. Unlike (2c) it is partitioned, contains lexical category labels and occurrences of Q.

What facts about event (1) does (2b) represent?<sup>14</sup>

It represents a stage between the formation of my intentions to use words in my memory—that is, the intentions represented by (2a)—and the formation of my phonetic intentions—that is, the intentions represented by (2c). Unlike (2a) and (2c), it does not represent intentions at all, though it does represent a mental set of sorts.

Remember events (1) and (5), the actual utterances? In the earlier one I pronounced the verb one way and in the later one I pronounced the same verb very

<sup>12</sup> See p. 23.

<sup>13</sup> For a different perspective, not committed to such an instrumentalist conception of phonological signs, see Bromberger and Halle (1997). In that paper—where we focus more closely on the semantics of phonological symbols—we assume that, WITHIN ANY THEORETICAL APPROACH TO PHONOLOGY, each phonological symbol must have the same semantic value wherever it occurs. We also explore there some consequences of that assumption in so far as it applies to symbols occurring in different lines of derivations and in different cells of OT tableaux.

<sup>14</sup> Some things we have said so far could be put in Austinian terminology (as in Austin 1975). (2a) modelled the formation of a PHATIC intention—that is, the intention to produce what Austin called a phatic act, 'the uttering of certain vocables or words ... belonging to, and as belonging to, a certain Vocabulary, conforming to and as conforming to a certain grammar'. (We leave out 'i.e. noises of certain types' as misleading.) (2c) modelled a PHONETIC intention—that is, the intention to produce what Austin called a phonetic act, 'the act of merely uttering certain noises'. (The 'merely' is unfortunate!) (2b) then models a stage in a MENTAL PROCESS through which the phatic mental set (or intention) gets transformed into the phonetic mental set (or intention).

differently. The facts underlying the difference can be surmised from a vast body of evidence, though they also happen to coincide with common beliefs. When I acquired my idiolect of English, I not only acquired words and other morphemes, but I also acquired rules. In producing these utterances I applied appropriate rules, and this led to different pronunciations of the same verb.

(2b) stands for a stage in the application of these rules.

We can even tell what stage.

As noted before, the verb *sell* and the noun *shelf* appear in two distinct guises in different utterances. Specifically, the verb *sell* undergoes a vowel change when combined with the past tense (a morpheme in my memory), and the noun *shelf* undergoes a change of the final consonant, when combined with the plural (also a morpheme in my memory). In other words, in producing these utterances, I invoked something like the rules in (8)

- (8) (a) Before [Q, Past] the stem vowel is [o] in the verbs *sell*, *tell*, ...  
 (b) Before [Q, Pl] the stem-final consonant is [+voice] in the nouns *house*, *knife*, *life*, *wife*, *shelf*, *mouth*, ...

(2b) then represents a stage after the application of rules (8).

The rules in (8) are not the only rules I applied to produce (1). I also applied rules to pronounce the morphemes represented by Q in (2a). Halle (1990) has argued that the relevant rules are statable roughly as follows:

- (9) Q → /n/ in env. X\_\_ Plural if X is *child*, *ox*, ...  
 /im/ in env. Y\_\_ Plural if Y is *kibbutz*, *hasid*, ...  
 /i/ in env. Z\_\_ Plural if Z is *alumn-*, *radi-*, ...  
 /ta/ in env. U\_\_ Plural if U is *stigma*, *schema*, ...  
 ø in env. V\_\_ Plural if V is *mouse*, *moose*, ...  
 /z/ in env. \_\_ Plural
- (10) Q → ø in env. X\_\_ Past if X is *sing*, *write*, ...  
 /t/ in env. Y\_\_ Past if Y is *buy*, *dream*, *mean*, ...  
 /d/ in env. \_\_ Past

So (2b) also represents a stage before the application of those rules!

But what do the phonetic symbols in (2b) represent? For that matter, what do they represent in the statement of the rules? The answer here is as before. They play a role as symbols in the formal computations of the theory. We conjecture that they also stand for specific aspects of the production of (1), but what that comes to exactly is not something clearly understood at this time and that must wait for further understanding of the nature of linguistic capacities and their actualizations.<sup>15</sup>

This double role assigned to phonetic symbols, we should point out, has a shortcoming: it slights certain important phenomena. So, for instance, it does not

<sup>15</sup> But see n. 13.

show that between the formation of (2a)'s referent and (2c)'s referent a kind of transubstantiation occurred through which mnemonic elements were converted into articulatory ones.

#### 4. ON THE NATURE OF TYPES

So far we have concentrated on a single and unique event, the utterance(1). We have done this because we hold that phonological theory, in so far as it purports to advance knowledge at all, is about such events and about the mental conditions and capacities responsible for their occurrence. Those are the sorts of things to which it is ontologically committed, or, as some followers of Quine would put it, those are the kinds of things over which it quantifies.

Our position may strike some as *prima facie* implausible, as simply conflicting with too many practices of phonologists.

Thus phonologists never mention or try to explain unique events like (1). Their papers and texts mention words, phrases, sentences, phonemes, that is, types, abstract entities outside time and space, devoid of causal histories and causal consequences. They do not mention utterances, events, or mental states. And, though phonologists do sometimes elicit tokens from themselves or from informants, they seem to do so only by way of getting facts about types. Thus their conclusions normally abstract from everything peculiar to such tokens, and, in so far as they seem to mention anything, they seem to mention types. In fact, no phonologists would probably interpret (2) as about (1), the utterance I produced umpteen minutes ago. How could they? How many have ever heard of that utterance! They would implicitly<sup>16</sup> take (2) as about a type possibly attested by something like (1) but attestable as well by other tokens—for instance, by the one that I now produce

(11) *The merchant sold shelves.*

And that last token, as Leibniz's indiscernability of identicals tells us, is not only numerically distinct from (1), since it occurred at a different time, but is also numerically distinct from its type, which does not occur in time at all.

Furthermore, phonologists, like all grammarians, strive for theories that neither undergenerate nor overgenerate, theories that predict some items and exclude others. But the relevant items cannot be tokens! If they were, any phonological theory, no matter how absurd, could always be trivially confirmed. Imagine, for instance, a theory that predicts that the following is in my language:

<sup>16</sup> We say 'implicitly' because, most phonologists explicitly avoid saying that they are theorizing about actual utterances by saying that they are investigating competence and not performance, as if that were sufficient to clarify the ontological nature of the entities they mention in practice.

(12) *The phonotofu sell yesterday many shelf.*

I would have confirmed that theory simply by having produced the token (12)! What is more, every theory of any interest would be demonstrably false since it would predict an infinite number of tokens for each person, whereas the total number of tokens is bound to be finite. Life is short! Even the most loquacious of us, no matter how long they live, will shut up forever at some point!

That may all be true; none the less, we do not believe that there are types! And so phonology cannot be about types. We admit (on empirical grounds) internalized grammars, but those exist as mental attributes of concrete, specific human individuals. We admit (on empirical grounds) internalized vocabularies, but those too exist as mental attributes of concrete, specific human individuals; and we admit token events (again on empirical grounds) but those are spatiotemporally located concrete events like (1). But types, as a kind of individual, as a kind of entity, we think, belong, with useful fictions like virtual optical images and stellar constellations, in the null class.

We cannot prove that there are no types. The notion is surely not self-contradictory, or even incoherent. Bromberger (1989) has argued that it is a coherent and even useful one. We just do not see any reason to think that any entities correspond to that notion. And we do not believe that phonology provides any evidence for such entities, or must presuppose their existence.

On the other hand, we do believe that phonology provides overwhelming evidence that tokens cluster into scientifically significant similarity classes. That does not imply that there are types besides tokens (not even as sets, or mereological sums of tokens, though such sets and sums may well exist). But it is sufficient to justify most of the practices we have mentioned, to make sense of the demand that theories should neither overgenerate nor undergenerate, and perhaps to explain the source of the illusion that there are types.

We will now explain that position in more detail.

Instead of producing (1) when I did, I could have produced a very different token. In fact, to fix ideas, here I go:

(13) *Two elephants study in Uppsala.*

And we could now go on and produce a derivation analogous to (2) for this new token. It would be a different derivation from (2). The last line, the first line, the intervening lines, the rules invoked, the items said to represent morphemes retrieved from lexical memory, all would be different. However, the two derivations would have one crucial thing in common: they would incorporate answers to the *same* questions. Different answers, but the same questions.

In other words, event (1) was open to the following questions. What morphemes in memory were intended? What was the representation of these morphemes in memory? What articulatory gestures were intended? What rules were invoked in the course of the formation of these intentions? and so on. (2) provides the answers

to these questions.<sup>17</sup> The last token (13) was open to the same questions. Its derivation, spelled out explicitly, would also provide answers to those questions, but the answers would be different. As we just said, different answers; the same questions. In fact, all spoken token events are open to these questions. Some warrant exactly the same answers. Token event (1) and the token event (11) do. Others do not warrant the same answers. Token events (1) and (13) do not. Token events that warrant the same answers are the ones we classify as BEING OF THE SAME TYPE. Those that do not, we classify as BEING OF DIFFERENT TYPES. Neither classification presupposes the existence of types. They presuppose only the possibility of certain similarities and differences among tokens.

That is all there is to talk of types, as far as we are concerned. But that is quite a lot, as we will now show.

Note, for instance, that each token warrants specific answers to these questions. We are unmitigated realists about this. We take the fact that each token warrants the specific answers it does to each of these questions to constitute truths about the world, not some artefact of our way of looking at things. It is a truth about the world that the answer to 'what was the first intended articulation underlying the production of (1)?' is 'vibration of the vocal cord (that is, [+voice]), constriction of the mouth cavity partially open (that is, [+continuant]), and so on', just as it is a truth about the world that the answer to 'How much does Sylvain Bromberger weigh?' is '175 lbs.'

Note furthermore that we also take it as a truth about the world—and a very different sort of truth—but not an artefact of our way of looking at things, that (1) has the property of warranting such answers at all. We might put this somewhat more technically. It is a truth about the world that event (1) had the DETERMINABLE property of having intended morphemes. And it is a truth about the world that each spoken token also does. Other events, even events with acoustic properties, do not have that determinable property. Noises made by our coffee pot, or coughs for instance, do not have it. That fact is of the same order as the fact that swinging pendula have periods, while standing rocks do not; that positive numbers have square roots, while trees do not; that the manuscript from which we are reading has a certain gravitational mass, while the ideas we are expressing do not. Determinable properties, by the way, such as period, square root, weight, and so on, are a kind of property presupposed by what-questions such as 'What is the period of ...?', 'What is the square root of ...?', 'What is the weight of ...?' Objects that do not have the property warrant no answer to the corresponding question. Those that do, and hence that have DETERMINATE instances of these determinable properties, warrant specific answers.

<sup>17</sup> Strictly speaking, to provide these answers it would have to be supplemented with references to the rules, as it was in the course of our discussion.

(1), (5), (13), and other tokens are, of course, open to many questions besides the ones answered in derivations such as (2). They warrant answers, for instance, to 'At what time was it uttered?', 'Where was it uttered?' If we were to use those questions to compare tokens as to sameness of type, we would end up with very different typological clusters. But we do not use those. We use questions that define the field of phonology. In other words, we use the questions that make phonology AS A THEORETICAL FIELD possible. That there are such questions, by the way, is also an empirical truth about the world!

An analogy that we have used elsewhere, taken from elementary chemistry, may be helpful here. Think for a moment about a sample of water, a real sample that you have 'experienced', as they say in California. That sample, like our tokens, is open to a number of questions—'Where was it situated when you experienced it?', 'Who owned it?', 'What did you do with it?'—that are of no scientific interest. But it is open to others that are of scientific interest, such as 'What is its boiling point?', 'What is its freezing point?', 'What is its molecular weight?' Other samples of stuff get the same answers to these last questions. They comprise all and only samples of water. Still other samples, though open to the same questions, get different answers. They comprise all and only the samples of other distinct substances—for instance, samples of gold all share one set of answers, samples of mercury share another set of answers, samples of sulphuric acid share yet another set, and so on. And still other samples of stuff do not hold answers at all to these questions. Pieces of sausage, for instance, or handfuls of mud, or buckets of shoelaces. These do not make up samples of a substance at all! The scientifically interesting questions collect bits of stuff into samples of SUBSTANCES. Not all bits of stuff. Some bits of stuff.

But what makes these questions scientifically interesting?

That should be obvious: the fact that their answers (when formulated in an appropriate notation) conform to lawlike, or at least computable, relationships. There are lawlike, or at least computable, relationships between boiling points, freezing points, and molecular weights. A theory can, therefore, be constructed on the back of these questions. And a certain attitude can be acquired. For instance, that these samples constitute a NATURAL domain, a domain that includes some (water, gold, mercury, and so on) but excludes others (sausage, mud, shoelaces); and that these samples have features that demand similar explanations. Of course, these attitudes are justified only if certain facts obtain—that is, only if certain lawlike relationships actually hold. For all we once knew, the world might have been otherwise.

A similar story, we believe, applies to utterances. Each utterance is open to a multiplicity of questions. The scientifically interesting ones are those whose answers across tokens stand in lawlike, computable relationships to each other. As we noted before, that there are such questions, if there are, is a fact about the world. An important and somewhat surprising fact. We believe there are such



facts.<sup>18</sup> If we did not we would not spend time on linguistics. And we believe that the questions we characterized as defining the domain of phonology are among these questions. If we did not we would not pursue phonology as we do. All of that then is implicated in our talk about types. But none of it entails that there are types as well as tokens.

We want to stress one crucial further fact about questions of scientific interest. They are not all given, they are not part of ordinary common sense, they must be smoked out, discovered, and their discovery can be an achievement more revolutionary than the discovery of a new phenomenon. Newton, for instance, discovered many things about the world, but his most important discovery (outside optics) would not have been possible without the discovery that physical bodies have MASS. That was the discovery of a new kind of question (and determinable property), revolutionary because answers to it turned out to stand in marvellous computable relation to other questions (for instance, 'What force is required to accelerate such and such object one foot per second?'). Aristotle, a very good scientist, did not have the concept of mass, and therefore could not have asked what the mass of the moon was, could not even know that he did not know what the mass of the moon was, and, of course, could not have fathomed that the answer to that question was related in systematic ways to the answer to other questions about the moon. Celestial mechanics was beyond even his imagination!

The questions that make phonology possible also had to be discovered. And their discoverers are the heroes of our field: Panini, Rask, Bopp, Saussure, Jakobson, Chomsky. Without these theorists, doing phonology would still consist in pedantically collecting odd curiosities. But the work of discovering the right questions is far from finished!

As we mentioned at the very outset, utterances form a natural domain with other noises, the domain of acoustical theory. To deny this would be like holding that elephants, because they have a sex life, are not, like rocks, physical objects subject to the laws of mechanics! On the other hand, to deny that utterances constitute an autonomous domain, the domain of phonology, would be like holding that, because elephants are subject to the laws of mechanics, like rocks, they have no sex life! And to deny a priori that there are systematic relationships between these two domains would be like denying a priori that there are systematic relationships between the mass of elephants and the character of their sex life.

We can now tell what we make of the fact that (2), though about (1), contained no information about time, speaker, and so on, and could have served for (11) as well. (2) contains only answers to questions of interest from the point of view of theoretical phonology. It could serve for any token that holds the same answers to

<sup>18</sup> The lawlike computable relationships are those that govern the production of utterances. Not of all utterances, but only of utterances produced by invoking rules and a lexicon. People can, of course, produce utterances without such invocation—when they produce gibberish, for instance. The crucial fact, not revealed by simple common sense, is that they can produce utterances that do invoke them.

those questions. Thus nothing about the abstract character of (2) entails that there be types besides tokens. Talk of types then is just a *façon de parler*.<sup>19</sup>

But that still leaves us with the requirement that phonological theory should not overgenerate or undergenerate. How do we construe the prohibition against overgeneration? Realistically: no phonological theory can be true of the world that generates derivations (combinations of answers to questions, like (2)) to which no token in our language (that is, produced as (1) was produced) can conform.<sup>20</sup> Phonological theories may not transcend the limitations on the production of tokens imposed by internalized rules and Vocabulary of the speakers. There is still no need to assume types.

What of the prohibition against undergeneration? We construe that also realistically: no phonological theory can be true of the world that cannot generate derivations to which tokens in our language (that is, produced as (1) was produced) could conform. Phonological theories may not exceed the limitations on the production of tokens imposed by internalized rules and Vocabulary of the speakers. There is again no need to assume types.

Our construal of these principles is stated with the aid of modalities (expressed by *could*) that vex many semanticists. But that is no reason to admit types. We find admission of types unenlightening as substitutes for these modalities, and at least as vexing.

Two final comments.

Some people may object to our way of looking at phonology on the grounds that it construes phonology as about performance and not about competence.<sup>21</sup> If they mean that we view phonology as about processes in real time responsible for the occurrence of tokens, they are right about our view. But we do not see this as an objection. If they mean that we view phonology as having to take into account contingencies of production over and above those traceable to knowledge of language, then they misconstrue our view. We do not.

Some will object that we have loaded phonology with unwarranted assumptions. Do speakers REALLY retrieve morphemes from their memory, invoke rules, go through all these labours when speaking? We think they do. In fact, we would like to know more about how they do it. We may be mistaken. Time will tell. But intuition will not. Clearly speakers are not aware of performing such actions. But then we perform many actions like zombies (to borrow a phrase from Ned Block). That is how we learn language, recognize faces, and solve most of our problems.

Some will object that our outlook leaves out entirely that tokens are not only

<sup>19</sup> We ourselves resort to this *façon de parler* even in this chapter, when, for instance, we speak of morphemes, phonemes, etc.

<sup>20</sup> Admittedly this answer requires elaboration. For instance, it seems to avoid commitment to phonological types at the price of commitment to types at the metalinguistic level—that is, at the level of the notation of the theory. We think that this appearance can be dispelled, but to do so would require a long discussion of theoretical formalisms. In any case, we are not claiming that there are no abstract entities at all. We are claiming that phonology is not about types.

<sup>21</sup> See e.g. Lindsey and Harris (1990).

uttered but are also recognized. That is indeed a big hole in our account so far. But it calls for another paper, not abstract types.<sup>22</sup>

Let us then return to the topic of our title, the ontology of phonology. What must the 'furniture of the world' include if phonological theory, as we conceive it, is to have a chance of becoming true of that world? It is a long list: agents with phonetic intentions, tokens, mind/brains with vocabularies and rules, articulators, and so on, in complicated interrelations. It does not have to include types. And if, perchance, the world does include types, phonology has nothing to say about them. But then, probably no branch of linguistics does.

## REFERENCES

- Austin, J. L. (1975), *How to Do Things with Words*, 2nd edn., ed. J. O. Urmson and M. Sbisà (Cambridge, Mass.: Harvard University Press).
- Bromberger, S. (1989), 'Types and Tokens in Linguistics', in A. George (ed.), *Reflections on Chomsky* (Oxford: Basil Blackwell), 58–89. Also in Bromberger (1992).
- (1992), *On What We Know We Don't Know* (Chicago: University of Chicago Press, CSLI Publications), 170–208.
- and Halle, M. (1986), 'On the Relationship between Phonetics and Phonology', in J. S. Perkell and D. H. Klatt (eds.), *Invariance and Variability in Speech Process* (Hillsdale, NJ: Lawrence Erlbaum), 510–20.
- (1989), 'Why Phonology is Different', *Linguistic Inquiry*, 20/1: 51–70.
- (1994), 'The Ontology of Phonology', in D. Prawitz, B. Skyrms, and D. Westerståhl (eds.), *Logic, Methodology and Philosophy IX* (Amsterdam: Elsevier), 725–743.
- (1997), 'The Contents of Phonological Signs: A Comparison between their Use in Derivational Theories and in Optimality Theories', in I. Roca (ed.), *Derivations and Constraints in Phonology* (Oxford: Oxford University Press), 93–122.
- Burton-Roberts, N. (1994), 'Ambiguity, Sentence and Utterance: a Representational Approach', *Transactions of the Philological Society*, 92: 179–212.
- (this volume), 'Where and What is Phonology? A Representational Perspective'.
- Cartwright, R. (1987), *Philosophical Essays* (Cambridge, Mass.: MIT Press), 33–53.
- Chomsky, N. (1986), *Knowledge of Language* (New York: Praeger).
- Halle, M. (1990), 'An Approach to Morphology', in J. Carter, R. Dechaine, W. Philip, and T. Sherer (eds.), in *Proceedings of North East Linguistic Society* (Armherst, Mass.: GSLA, Department of Linguistics, University of Massachusetts), 1: 150–84.
- (1997), 'Some consequences of the representation of words in memory', *Lingua*, 100: 91–100.
- Hutton, C. M. (1990), *Abstraction and Instance—the Type-Token Relations in Linguistic Theory* (Oxford: Pergamon Press).
- Katz, J. J. (1990), *The Metaphysics of Meaning* (Cambridge, Mass.: MIT Press).
- Lenneberg, E. H. (1967), *Biological Foundations of Language* (New York: Wiley).
- Liberman, A. M. (1996), *Speech: A Special Code* (Cambridge, Mass.: MIT Press).
- and Mattingly, I. G. (1985), 'The Motor Theory of Speech Revised', *Cognition* 21: 1–36.
- (1989), 'Specialization for Speech Perception', *Science* 243: 489–94.
- Lindsey, G. and Harris, J. (1990), 'Phonetic Interpretation in Generative Grammar', *UCL Working Papers in Linguistics*, 2: 355–69.
- Peirce, C. S. (1933), *Collected Papers of Charles Sanders Peirce*, iv, ed. C. Hartshorn and P. Weiss (Cambridge, Mass.: Harvard University Press).
- Quine, W. V. (1987), *Quiddities* (Cambridge: Harvard University Press), 216–19.
- Strawson, P. F. (1950), 'On Referring', *Mind* 235: 320–44.
- Szabo, Z. (1999), 'Expressions and their Representations', *Philosophical Quarterly*, 49: 145–63.
- Wetzel, L. (1993), 'What Are Occurrences of Expressions?', *Journal of Philosophical Logic*, 22: 215–20.

<sup>22</sup> But on the relationship between production and recognition see Halle (1997) and Liberman (1996).