1 Introduction

This chapter provides an overview of how experimental work has informed phonological theories, and vice versa. This chapter starts with a historical overview; when phonology was being established as its own area of research, there was a sharp division between phonetics and phonology. This division was called into question, and issues concerning the phonetics–phonology interface are currently being extensively pursued by an approach that is now known as Laboratory Phonology. After this historical overview, I discuss in some detail how phonetic experiments and phonological theories have informed each other.

1.1 The tension between phonetics and phonology

When phonology was being established as its own area of research, it was often assumed that phonology and phonetics were independent of one another. For example, Trubetzkoy (1939: 11) stated:

The speech sounds . . . possess a large number of acoustic and articulatory properties. All of these are important for the phonetician since it is possible to answer correctly the question of how a specific sound is produced only if all of these properties are taken into consideration. Yet most of these properties are quite unimportant for the phonologist.

We still sometimes witness a sharp divide between phonetics and phonology in the current literature: some claim that phonology is an abstract, substance-free computational system, which should be separated out from phonetics: “patterns of phonetic substance are not relevant to phonological theory strictly defined” (Hale & Reiss 2000: 158; see Blaho 2008 for a recent review of this position). There has been an uneasiness about integrating phonetics into phonological studies, because of a belief that the phonetic module belongs to performance and the phonological module belongs to competence; i.e. phonetics does not belong to grammar per se (see e.g. Blaho 2008: 2). An assumption behind this claim is that phonetics involves automatic, universal mechanisms.
However, contrary to the view that phonetics consists of universal implementation rules, experiments have shown that phonetics is neither automatic nor universal; i.e. speakers control their phonetic behaviors, and cross-linguistic variation exists in the realm of phonetics (Keating 1985, 1988a; Kingston & Diehl 1994; see also §2.3.1 for more discussion). In this sense, any adequate model of grammar must integrate phonetics as a part of its model.

Phoneticians also responded to the thesis that we can and should study phonology without considering the phonetic mechanisms behind phonology (see Lindblom 1962, Ohala 1990b, and Diehl 1991 for general discussions). Perhaps the best-known advocate of the objection is Ohala (1990b, among many other references). His general point is that many phonological patterns can be explained in terms of articulatory and perceptual factors, and therefore purely phonological explanations without considering phonetic substances can be arbitrary, circular, and post hoc. For example, many languages lack voiced stops in their inventory, a situation for which we could propose a redundancy rule \([-\text{son}, -\text{cont}] \rightarrow [-\text{voice}]\), as in SPE (Chomsky & Halle 1968), or a constraint \(*[-\text{son}, -\text{cont}, +\text{voice}]\), but these approaches miss the aerodynamic reason behind the dispreference for voiced stops.

In order for speakers to maintain voicing, intraoral air pressure must be lower than subglottal air pressure, but the airflow required for voicing increases intraoral air pressure when the airway is significantly occluded. The increase in intraoral air pressure in turn makes it difficult to satisfy the aerodynamic condition. For this reason, it takes additional articulatory effort – e.g. larynx lowering, tongue advancement, etc. – to keep the intraoral air pressure sufficiently low to maintain voicing during stop closure (Jaeger 1978; Westbury 1979; Ohala 1983).

In addition to this kind of articulatory difficulty, perceptual factors demonstrably affect phonological patterns as well. For example, non-low back vowels are usually rounded, so that we could postulate a redundancy rule \([-\text{low}, +\text{back}] \rightarrow [+\text{round}]\) or a constraint \(*[-\text{low}, +\text{back}, -\text{round}]\), but again these explanations miss the generalization that rounding, by enlarging a resonance cavity, enhances the F2 difference between back and front vowels (Stevens et al. 1986; Diehl & Klunder 1989; Diehl 1991).

Finally, psycholinguistic factors also seem to play an important role in shaping phonological patterns. For instance, word-initial segments provide important cues for word recognition (Nooteboom 1981; Hawkins & Cutler 1988). Speakers thus seem to disfavor making phonological changes in word-initial positions, because such changes would result in difficulty in word recognition (Beckman 1997; Kawahara & Shinohara 2010).

In summary, there seem to be phonetic and/or psycholinguistic reasons behind many, if not all, phonological patterns. Therefore, according to Ohala, purging phonetic and psycholinguistic factors from phonological theory is misguided.

1.2 Current situation

Thus, on the one hand, there has been some reluctance to incorporate phonetic (and psycholinguistic) factors into phonological explanations. On the other hand, phonetics does seem to offer some insights into phonological patterns. For this

---

1 The tension between phonetics and phonology existed before the inception of generative phonology (Chomsky & Halle 1968). See Ohala (1990b, 1999) for reviews of the history.
(and perhaps other) reason(s), there has been some unfortunate intellectual tension between phonetics and phonology, which Ohala describes as a “turf war” (1990b: 168), where people from each discipline felt that they had to delineate and defend their own territory.2

The situation, however, has been changing, as we witness the rise of a general approach which has come to be known as “laboratory phonology” (see Cohn 2010 for sociological aspects of the development of laboratory phonology in the field of general linguistics). The following quote from Beckman and Kingston (1990: 5) succinctly summarizes the spirit of this approach:

We believe that the time has come to undo the assumed division of labor between phonologists and other speech scientists; we believe this division of labor creates a harmful illusion that we can compartmentalize phonological facts from phonetic facts. At the very least, we maintain that the endeavor of modeling the grammar and the physics of speech can only benefit from explicit argument on this point.

As the following discussion shows in more detail, many experimental studies have contributed to theoretical debates. The rest of the discussion proceeds as follows. In §2, I discuss how experimental approaches have informed phonological theories. In §3, I reverse direction and discuss cases in which theories have informed experiments. Although I try to be comprehensive in my review, there is necessarily a limit. For further examples and discussion, readers are referred to contributions in the Laboratory Phonology series (Kingston & Beckman 1990 et seq., as well as in other volumes and papers devoted to this issue (Ohala 1986b; Ohala & Jaeger 1986; Diehl 1991; Hayes et al. 2004; Kingston 2007; Solé et al. 2007; Coetzee et al. 2009).

2 How experiments have informed theory

2.1 Beyond introspection-based data

In generative linguistics, native speakers’ intuition – or introspection – is the primary source of data, because “the set of grammatical sentences cannot be identified with any particular corpus of utterances obtained by the linguist in his field work” (Chomsky 1957: 15). Since generative phonology aims to study competence, i.e. what speakers know about their language, rather than performance, i.e. how speakers use the language, the only way to assess competence, it was believed, was introspection (though see Schütze 1996 for a critical discussion). Contrary to this research tradition, phonetic and psycholinguistic experiments have offered important insights into knowledge of grammar.

2.1.1 Wug tests

The first good example of experiments that have complemented the introspection-based approach is a wug test. In wug tests, named after an experiment by Berko (1958), native speakers are asked to pronounce novel words. Berko tested

---

2 An anonymous reviewer points out that there may also be “punting,” when people say that some other subfield is responsible for a phenomenon that they cannot account for.
whether English-speaking children acquire a rule of voicing assimilation in the English plural and other suffixes, and found that, given a nonce word like wug, most children pluralize it as wug[z], not as *wug[s], showing that English-speaking children know that the plural suffix and the stem-final consonant must agree in voicing. In this way, the wug test has been used as a litmus test for the productivity of a phonological generalization. Recent years have witnessed a renewed interest in wug tests, which has provided some important insights into phonological knowledge, as summarized in (1).

(1) a. A standard assumption in generative phonology is that speakers assign a simple dichotomous grammatical/ungrammatical judgment to linguistic structures. In other words, speakers should treat all attested structures as equally good, and all ungrammatical structures as equally bad. However, several wug tests revealed that speakers can distinguish the relative grammaticality of two (un)grammatical structures (Shinohara 2004; Zuraw 2007).

b. More generally, the results of wug tests often show stochastic, rather than dichotomous, patterns (Albright & Hayes 2003; Hayes & Londe 2006).

c. Some experiments have shown that the probability of speakers applying a certain phonological process in a wug test reflects the frequency of the items that undergo that phonological process in their language (Bybee 1999; Zuraw 2000; Albright & Hayes 2003; Ernestus & Baayen 2003; Hayes & Londe 2006; Hayes et al. 2009; chapter 90: frequency effects).

d. In some experiments, speakers either fail to replicate some statistical patterns in the lexicon (Becker et al. 2008) or at least show bias against reproducing some arbitrary, though statistically significant, patterns in the lexicon (Hayes et al. 2009).

e. Some phonological patterns are not productive (at least under a wug test), which leads to the suspicion that they are not a part of the speakers’ grammar. Patterns whose productivity wug tests have failed to reveal include English velar softening (Ohala 1974; though see Pierrehumbert 2006), Japanese verb conjugations (Vance 1987: ch. 12) and Polish raising (Sanders 2001; see also Zimmer 1969 for a test of morpheme structure conditions in Turkish).

2.1.2 Well-formedness judgment studies

Another type of experiment which complements generative phonology’s introspection-based approach is well-formedness judgment experiments. In these experiments, native speakers are asked to judge the naturalness of particular words or phonological processes (they can also take the form of word-likeness judgments). These experiments, as with wug tests, reveal, for example, that speakers can distinguish the relative grammaticality of two (un)grammatical structures (Pertz & Bever 1975; Coetzee 2008) and show that grammatical patterns exhibit a stochastic, rather than a simple dichotomous grammatical/ungrammatical, distinction (Hayes 2000; Albright & Hayes 2003; Fanselow et al. 2006). Well-formedness judgments are also known to reflect the frequency of the target items (e.g. Frisch et al. 2000; see also chapter 90: frequency effects).
2.2 Addressing the quality of phonological data

2.2.1 Re-evaluating phonological data

Experiments have also re-evaluated what is phonological and what is not. Phonetic experiments have shown that many textbook examples of "phonological patterns" do not involve categorical changes but instead involve gradient changes, suggesting that they might be phonetic processes (chapter 89: Gradience and Categoricality in Phonological Theory). For example, English was thought to have a vowel nasalization rule before a nasal consonant, as in [bin] bean and [din] dean (Fromkin & Rodman 1998: 280–281). Cohn (1993), however, based on an instrumental study measuring patterns of nasal airflow, showed that English nasalization differs from contrastive nasalization in French in that, the closer to the nasal consonant, the more nasal airflow was detected within a nasalized vowel. English nasalization is therefore gradient rather than categorical, in the sense that it does not alter the whole segment but instead the degree of nasalization changes within a segment. For this reason, Cohn concluded that English nasalization belongs to phonetics. Many other examples of phonological patterns have been argued to show similar gradient properties, which I list in (2):3

(2) a. Arabic tongue backing (emphasis) spreading (Keating 1990 and references cited therein; see also chapter 25: Pharyngeals)
   b. English /l/-velarization in coda (Sproat & Fujimura 1993)
   c. English flapping (Fox & Terbeek 1977; de Jong 1998; chapter 113: Flapping in American English)
   e. English phrasal palatalization (Zsiga 1995)
   f. English and French schwa deletion (Fougeron & Steriade 1997; Davidson 2006b; chapter 68: Deletion)
   g. Japanese tonal spreading in unaccented words (Pierrehumbert & Beckman 1988)
   h. Russian vowel reduction in second pretonic syllable (Barnes 2002; but cf. Padgett & Tabain 2005; chapter 79: Reduction)

The abundance of such examples led Hayes to state “I occasionally wondered, ‘Where is the normal phonology that I was trained to study?’” (1995: 68).

The list in (2) shows that many patterns that were believed to be phonological have turned out to be phonetic. A more complex example comes from the domain of intonation. In Japanese and many other languages, the height of tones generally decline toward the end of an utterance. The question arose whether this pattern of declination is due to phonetics or phonology. One could posit that

3 Davidson (2006a) demonstrates that a “schwa” inserted in English speakers’ production of non-native clusters differs from a lexical schwa. She argues that this “schwa” results from gestural mis-coordination, and hence differs from phonological epenthesis (see also Hall 2006 for related cross-linguistic phenomena). However, within the framework of Articulatory Phonology (Browman & Goldstein 1986), she also proposes that gestural mis-coordination arises in the phonological component, rather than in the phonetic component.
this declination is phonetic (Fujiwaki & Sudo 1971; see Poser 1984: 200, for more references); for example, subglottal air pressure decreases towards the end of an utterance, and the height of tones naturally drops. On the other hand, McCawley (1968) proposes a phonological rule in Japanese that changes a high tone to a mid-tone after another high tone within a phrase. It turns out that it would be most fruitful to approach intonation from both perspectives. Poser (1984) argues that Japanese has both local lowering of H after another H(L) – which seems phonological – as well as gradient, steady declination throughout the utterance, which is phonetic. Beckman and Pierrehumbert argue that a similar hybrid approach accounts for the complex pattern of intonation in both English and Japanese (Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988).

In addition to helping us to decide whether patterns under discussion are phonological or phonetic, some studies have called into question the existence of some phonological patterns per se. Based on the traditional description of Tswana, Hyman (2001) discusses a case of post-nasal devoicing, but Gouskova et al. (2006) argue, on the basis of a production experiment, that Tswana may not have a process of post-nasal devoicing after all. A later study, however, suggests that some, though not all, speakers do show evidence for post-nasal devoicing (Coetzee et al. 2007). A general lesson we can draw from this series of studies is that careful instrumental experiments help us to establish whether phonological patterns under discussion really exist.

2.2.2 **Incomplete neutralization**

While it is standardly assumed that phonological processes involve categorical changes (see §2.2.1 and §2.3.2), some experiments have called this assumption into question. Port and O’Dell (1985) report a production experiment on German where they found some acoustic differences between underlying voiceless stops and “voiceless” stops that are underlyingly voiced but devoiced by coda devoicing (Chapter 69: Final Devoicing and Final Laryngeal Neutralization). They found appreciable differences between these two categories in terms of preceding vowel duration, closure voicing duration, closure duration, and aspiration duration. Further, they demonstrated that listeners can detect the differences between the two categories at more than chance frequency. They argue that coda devoicing in German is therefore incomplete.

Subsequent studies have found other cases of incomplete neutralization in many languages, including Cantonese (Yu 2007), Catalan (Dinnsen & Charles-Luce 1984), Dutch (Ernestus & Baayen 2006, 2007; Warner et al. 2006), English (Fourakis & Port 1986; Ohala 1986a), Japanese (Mori 2002), Lebanese Arabic (Gouskova & Hall 2009), Polish (Slowiaczek & Dinnsen 1985; Slowiaczek & Szymanska 1989) and Russian (Chen 1970; Dmitrieva 2005; Padgett & Tabain 2005; see also Chapter 80: Mergers and Neutralization). Some studies have argued, however, that these experimental results are largely or entirely due to extra-grammatical factors such as speakers’ familiarity with English, orthographic influences and hyper-articulation in a laboratory setting (Fourakis & Iverson 1984; Jassem & Richter 1989; Warner et al. 2006). The status of incomplete neutralization is much debated in the literature (see Warner et al. 2004; Port & Leary 2005 for reviews), but these

---

4 Downstep may apply iteratively (Poser 1984; Kubozono 1988), resulting in quasi-gradient behavior. See §2.2.1 and §2.3.2 for the categorical nature of phonological patterns.
experiments at least show that we need to be careful when talking about the
categoricality of phonological changes. See §2.3.2 for more on the discussion
on the categorical nature of phonological alternations.

2.3 Bearing on the architecture of the grammar

Not only have experiments served to evaluate the quality of phonological data,
some phonetic studies have provided important insights into the general archi-
tecture of the grammar.6

2.3.1 Against universal phonetics

In SPE, the output of phonology was considered to be “the phonetic transcription”
(SPE: 293), which lacked “properties of the signal that are supplied by universal
rules” (SPE: 235). Keating (1985, 1988a) characterizes this view as phonetics
involving universal, automatic rules (see also Kingston & Diehl 1994: §1.2).
Phonetic studies soon showed that this view is too simplistic. For example, Chen
(1970) compared durations of vowels before voiced consonants and those before
voiceless consonants in seven languages (English, French, Russian, Korean, German,
Spanish and Norwegian), and showed that different languages show different
degrees of lengthening before voiced consonants. Keating (1979) (reported in
Keating 1985) followed up on this result, and showed that neither Czech nor Polish
shows a reliable effect of voicing on preceding vowel duration. It therefore seems
that the degree of lengthening before voiced consonants is language-specific.
Similarly, an acoustic experiment by Port et al. (1980) showed that in Japanese
vowel durations are heavily affected by the duration of adjacent consonants,
but in Arabic such patterns are not evident, concluding that rhythmic compen-
sation is not universal. These examples show that phonetic implementation is
neither automatic nor universal. See Port and Leary (2005) for recent summaries
of language-specific phonetic patterns.

2.3.2 The phonetics–phonology divide

As briefly discussed in §2.2.1, many experiments have identified a crucial differ-
ence between phonetics and phonology: phonological patterns involve complete
categorical changes, whereas phonetics yield gradient outcomes (Keating 1990;
Cohn 1993, 2006; Zsiga 1995; Tsuchida 1997; Barnes 2002). Experimental results
played an essential role in establishing this difference. For instance, an electro-
palatographic study by Zsiga (1995) showed that English possesses two kinds of
palatalization: complete palatalization, which we find in a morphophonological
process, as in press [prɛs] vs. pressure [preʃər], and gradient palatalization, which
we find across a word boundary, as in miss you [mɪʃju]. Zsiga found that the

5 Some phonologists admit that some phonological changes are incomplete and propose a model
of phonology that handles incomplete neutralization (van Oostendorp 2008; Gouskova & Hall 2009). Others
consider the results of neutralization as lacking phonological/phonetic specifications (Steriade 1995,
1997), following the theory of phonetic underspecification (Keating 1988b; also Hsu 1996, cited in Steriade
1995, 1997). Yet others consider these incomplete neutralization patterns to be implemented in the
phonetic component (Fourakis & Port 1986).

6 Another topic that would fit in this subsection is the search for the phonetic basis of distinctive
features. Due to space limitations, I cannot provide comprehensive discussion. See Kingston (2007)
for a summary.
former [f] is [f] throughout its constriction, whereas the latter [f] starts like an [s] and ends like an [ʃ]. An explanation we can give is that the former process involves a categorical phonological change, whereas the latter process is a gradient phonetic gestural overlap.

Pycha (2009) demonstrated another difference: comparing phonological lengthening (i.e. gemination) and phonetic lengthening at phrase edges in Hungarian, she found that phonological lengthening always targets the closure phase of affricates, whereas phonetic phrase edge lengthening affects portions that are adjacent to the boundaries. In this way, experiments have identified characteristics of phonetics that distinguish it from phonology. See Keating (1996: 263) for constellations of other properties that distinguish phonetics and phonology. See also Anderson (1981) for general discussion on the phonetics–phonology divide.

2.3.3 The phonetics–morphology interface?

As exemplified by the two palatalization processes in English, morphophonological processes tend to involve categorical changes whereas phonetic processes yield gradient outputs. A general assumption in generative studies is thus that phonology can be sensitive to morphology, but phonetics is not. The inaccessibility of morphological structures to phonetics was assumed in Chomsky and Halle (1968), where morphological boundaries are erased at the end of each transformational cycle (SPE: 15). The Bracket Erasure Convention in Lexical Phonology (Kiparsky 1982) also removes morphological boundaries after each level of derivation (chapter 85: cyclicity). As a result, word-internal structures are inaccessible to later post-lexical rules or phonetics.

As an illustration, take the case of minimal word requirements. Many languages require (lexical) words to be of certain minimal length, and this requirement is expressed in terms of abstract prosodic units (McCarthy & Prince 1986), but not in terms of raw phonetic duration (Cohn 1998). For example, even though English tense [i] is shorter than lax [æ] in raw duration ([i] = 100 milliseconds; [æ] = 123 milliseconds, according to Strange et al. 2004), [pi] is well formed, but [pæ] is not. Therefore, the minimal word requirement operates on abstract phonological units rather than on raw phonetic duration. This sort of requirement can be sensitive to morphological information. For example, in Yoruba, only nouns are required to be maximally disyllabic (Pulleyblank 1988: 250, fn. 24). On the other hand, no known languages seem to vary raw phonetic durations depending on morphological categories. Phonological requirements, therefore, may refer to morphological information whereas phonetic implementation cannot. This thesis has been taken for granted and rarely questioned or addressed in the phonological literature.

However, Cho (2001) directly addressed this issue using EMA (electromagnetic articulography), and found that in Korean gestural timing is more variable across a morpheme boundary than within a morpheme and also more variable across a non-lexicalized compound boundary than across a lexicalized compound boundary. Also, Sproat and Fujimura (1993) used X-ray microbeam technology and compared the amount of dorsal retraction of English coda [l] at various boundaries.

7 In turn, morphological processes can be sensitive to phonological information (i.e. phonologically conditioned allomorphy; see McCarthy 2002: 183 for references), but do not seem to be controlled by phonetic information (though see Bybee 1999 for a case of a morphological pattern that manipulates a non-contrastive feature).
including Level I and Level II boundaries, and found a difference between these two contexts. These experimental findings suggest that morphological boundaries may be visible to phonetic implementation rules.\(^8\)

Another debate was initiated by Steriade (2000), who challenges the immunity of phonetics to morphological information. She argues that there are cases of phonetic analogy, a requirement that paradigmatically related words be phonetically similar. For example, derived words are required to be identical in raw phonetic duration to their corresponding bases. This phonetic analogy is proposed to explain why flapping applies in words like capi\(\text{t}al\)istic (cf. capi\(\text{t}\)al), whereas it fails to apply in words like militaristic (cf. mili\(\text{t}\)ary) with a similar stress pattern: the applicability of flapping in derived words depends on whether flapping is possible in the base words. However, Riehl (2003) found in a production experiment that the transfer of flapping from a base form to related words was not robustly observed. Based on an acoustic study Riehl also challenged the assumption that the distinction between \[t\] and \[r\] is made solely in terms of constriction duration (see Fox & Terbeek 1977 and de Jong 1998 for the phonetics of flapping; also chapter \(113\): flapping in American English).

In relation to incomplete neutralization discussed in §2.2.2, Ernestus and Baayen (2006, 2007) argue for another case of morphological influence on phonetics. They found that there is slight voicing left in “devoiced” final consonants, and argue that this voicing is due to the activation of morphologically related words with a voiced consonant (chapter \(83\): paradigms). In summary, whether phonetics has access to morphological information or not is still under debate; experiments will be able to contribute much to this debate (see Bybee 1999, Barnes & Kavitskaya 2003, Davis 2005, Cohn 2006, and Yu 2007 for further discussion on the phonetics–morphology interaction and phonetic analogy).

### 2.4 Arguments for and against the psychological reality of grammar

Not only have experiments addressed what grammar should look like, they have also examined whether grammar is psychologically real in speakers’ minds. Many experiments have addressed the question of whether the rules, constraints, and structures that linguists posit are psycholinguistically real or are merely theoretical devices that help us explain the linguistic patterns (Zimmer 1969; Ohala 1974, 1986a; Cena 1978; McCawley 1986). A general concern behind this work is that the psychological reality of a grammatical postulate was sometimes confused with the analytical success of that postulate. As McCawley (1986: 28) puts it, “Chomsky’s ([1986]) policy that the subject matter of linguistics is psychological in nature does not provide any reason for assuming that the purported facts that linguists have hitherto adduced as evidence for or against particular analyses are psychological in nature, nor even that they are strictly speaking facts.” Psychological reality of phonological data should not be taken for granted, and must be explicitly tested. Some wug tests in fact revealed that some phonological patterns are not reflected in speakers’ behaviors (see (1e)).

---

\(^8\) A question that remains with respect to these results is whether the differences could be attributed to differences in the presence of prosodic boundaries like foot boundaries or prosodic words, which the phonetics is presumably able to see.
Recent development of experimental techniques has allowed us to address the question of psychological reality from a different perspective. In particular, a number of perception experiments showed that phonological constraints affect speech perception—given ambiguous acoustic signals, speakers are biased against categorizing the stimuli as those not allowed by their phonological grammar (Massaro & Cohen 1983; Pitt 1998; Dupoux et al. 1999; Moreton 2002; Berent et al. 2007; Coetze 2008). A classic work by Massaro and Cohen took advantage of word-initial phonotactic restrictions in English, where only [P] is allowed after [t], only [l] is allowed after [s], and both are allowed after [p]. They created a continuum from [z] to [l] by varying F3 and presented the continuum in these contexts, and found that speakers judge tokens as [z] most frequently after [t], less frequently after [p], and least frequently after [s]. These results showed that phonotactic restrictions in speakers’ grammars affect how they categorize the speech signals.

Extending this work, some studies showed that some particular phonological hypotheses are psychologically real. For example, in Japanese, only foreign words, but not native or Sino-Japanese words, allow word-final long [aa] and singleton [p] (Ito & Mester 1995). Moreton and Amano (1999) showed that once listeners hear [p] in the stimuli, cueing foreignness of the stimuli, then they are more likely to judge the word-final [a] as long [aa]. Gelbart and Kawahara (2007) extended this result, and showed that as long as real foreign words are presented, a similar bias towards allowing word-final long [aa] perceptually is observed, even in the absence of phonological cues to the lexical affiliation. See Gelbart (2005) for similar results from other languages.

On a slightly different line of research, acoustic studies have provided evidence for particular prosodic structures (Maddieson 1993; Broselow et al. 1997; Frazier 2006) or tonal representations (Morén & Zsiga 2006). Broselow et al. (1997) show that language-particular prosodic structures, each motivated in terms of stress placement, are manifested in different phonetic implementation patterns. Yet another line of research argues for the psychological reality of underspecification (Archangeli 1988). For example, a priming study by Lahiri and Reetz (2002) shows that labial and dorsal signals can activate coronal input. They argue that, assuming coronals are underspecified in the mental lexicon, all labial, coronal, and dorsal consonants can be matched up with underlying coronals (see also Lahiri & Marslen-Wilson 1991; Chapter 12: CORONALS). These studies aim to show that theoretical devices that have been proposed, such as lexical stratification, prosodic structure or underspecification, may not merely be abstract theoretical constructs, but may be psychologically real, influencing our speech behaviors (see Goldrick, forthcoming, for discussion).

2.5 Sources of phonological patterns

Finally, many experiments have addressed the issue of sources of phonological patterns. This tradition has been most rigorously pursued by Ohala (e.g. 1983, 1990b), but has been taken up by many other researchers. For example, in many languages [k] (or [k] before front vowels) changes diachronically into [ʃ]. The ubiquity of this sound change (and its synchronic correspondence) may be attributed to the acoustic affinity between [k] and [ʃ] (Ohala 1989; Guion 1998; Chang et al. 2001; Wilson 2006). Raising of F2 via palatalization makes [k] sound similar to palatal consonants, and a long period of aspiration of dorsal [k] makes
it sound similar to an affricate. A perception experiment by Guion (1998) demonstrated in fact that listeners often misperceive [k] as [f], showing that the sound change [k] → [f] may be due to the acoustic similarity. Furthermore, Chang et al. (2001) point out that the directionality of the [k] → [f] change is rarely if ever reversed, and demonstrate experimentally that listeners may perceive [k] as [f], but not vice versa.

Another example is provided by the fact that in many languages a vowel must be long after a glide. Traditionally, this restriction has been analyzed as a case of compensatory lengthening: the first vowel in vowel sequences obtains a mora by a universal convention, loses its mora when it becomes a glide, with the floating mora being reassocated with the following vowel, resulting in a long vowel (Hayes 1989; see also chapter 64: compensatory lengthening). Myers and Hansen (2005) offer an alternative explanation: given a sequence of two vocoids, the boundary between them is blurry, and listeners may misattribute the gradient transition to the second vowel. The misattribution would result in a percept of long second vowels. Their perception experiments supported their hypothesis: the longer the transition duration, the more likely listeners judge the second vowel to be long (chapter 20: the representation of vowel length).

The list of other experiments which have searched for the basis of phonological patterns would include, but is not limited to, the following: Kohler 1990; Ohala 1990a; Hura et al. 1992; Kawasaki-Fukumori 1992; Huang 2001; Hume & Johnson 2001; Barnes 2002; Mielke 2003; Kawahara 2006; Kochetov 2006; Myers & Hansen 2007) as well as those discussed in §3.1. See also Blevins (2004) and Ohala’s other work (e.g. Ohala & Lorentz 1977; Ohala 1981, 1983) for further cases of phonetic origins of phonological patterns.

3 Experiments informed by phonology

So far I have been focusing on how experiments have informed phonological theories. However, the communication is by no means one-way. So we now turn our attention to how phonological observations and theories helped us design phonetic experiments and led to important discoveries.

3.1 Experiments motivated by phonological observations

As discussed in §2.5, many experiments have attempted to make sense of why certain phonological patterns occur. Put in a different perspective, this tradition has allowed us to reveal aspects of our phonetic systems by addressing why phonology works in the way that it does. To illustrate this point with another example, an influential tradition of this line of research is that of Adaptive Dispersion Theory, initiated by Lindblom and his colleagues (Liljencrants & Lindblom 1972; Lindblom 1986; Diehl et al. 2004) and pursued by a number of studies (e.g. Flemming 1995; Boersma 1998; Padgett 2002). This theory sets out to address why languages have the sets of vowels that they have. For example, languages that have three contrastive vowels usually have [a i u] rather than, say, [a ï u], and languages that have five contrastive vowels have [a i u e ø] rather than [a ï u e x]. The general idea is that speakers keep contrasting elements maximally (Liljencrants...
3.2 Experiments motivated by phonological theories

Not only can phonological observations lead to interesting phonetic hypotheses and experiments, specific phonological hypotheses can sometimes provide a guideline for where to look in experimental work. For example, in traditional analyses of Japanese intonation (Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988; Venditti 2005), Japanese was not thought to have an Intonational Phrase. Selkirk (2005), however, based on cross-linguistic patterning, proposes a general theory of syntax–phonology mapping where clause edges should generally correspond to Intonational Phrase edges. Guided by this theory, Kawahara and Shinya (2008) investigate the intonational properties of clause edges in Japanese and find evidence for Intonational Phrase edges. In particular, they find that the left edges of clauses show larger initial rises and stronger pitch reset compared to VP edges, and clauses are also characterized by final lowering of tones, final creakiness and pause at their right edges. This work shows that theories can provide a guideline as to what to look for in phonetic studies.

Another example comes from articulatory studies on transparent segments in harmony contexts (Gick et al. 2006; Benus & Gafos 2007; Walker et al. 2008; chapter 91: vowel harmony: opaque and transparent vowels). Several works have proposed that autosegmental spreading is strictly local and can never skip a segment (Ní Chiosáin & Padgett 1997; Gafos 1998; Walker 1998). Transparent segments in harmony patterns pose a problem for this theory because it looks as though these segments are “skipped.” Recent articulatory studies showed, however, that “transparent” segments also undergo harmony (e.g. tongue body backing in back vowel harmony in Hungarian and the tip blade gesture in Kinyarwanda consonant harmony), without causing much perceptual effect. This outcome was as predicted by strict locality, because transparent segments, too, undergo harmony phonologically. Again, the theory of strict locality has led to experiments that reveal a non-trivial aspect of transparent segments in harmony contexts. See Hayes (1999) for related discussion on theory-driven experiments.

3.3 Testing specific phonological hypotheses and beyond

Specific phonological hypotheses can motivate specific hypothesis testing, which has often resulted in further insights into the intricacy of the phonetics–phonology interface. To take one example, Steriade (1997, 2009) proposes that the less perceptible a phonological contrast is (in a particular context), the more likely it is to be neutralized. Some work has shown that at least some contrasts that are likely to be neutralized are indeed less perceptible than non-neutralizing contrasts (Kawahara 2006; Kochetov 2006; chapter 80: mergers and neutralization). Kochetov (2006) showed further, however, that not all differences in phonetic
perceptibility are reflected in phonological patterns. Once again, Steriade’s specific hypothesis about the interaction between phonetic perceptibility and phonological patterns motivated experimental testing, which revealed the complex interaction between phonetics and phonology.

To take another example, many languages require that lexical words be minimally bimoraic or bisyllabic (§2.3.3). Japanese, however, allows monomoraic lexical words, and Mori (2002) tested whether Japanese does indeed violate the minimality requirement. She found that when monomoraic words are pronounced without a case particle, they undergo lengthening, while longer words do not show such lengthening. In this sense, Japanese does satisfy the minimal word requirement. However, she further found that lengthened monomoraic roots are not as long as bimoraic roots, instantiating a case of incomplete neutralization (§2.2.2).

To summarize, these experiments show that specific phonological hypotheses can inform experiments, which often in turn provide insight into the complex interaction between phonetics and phonology. The list of theories that has motivated specific experimentation includes the sonority sequencing principle (Broselow & Finer 1991), Optimality Theory’s (Prince & Smolensky 1993) transitivity of constraints (Guest et al. 2000), the Emergence of the Unmarked (Broselow et al. 1998), and positional faithfulness theory (Kawahara & Shinohara 2010).

4 Summary

Phonetic and psycholinguistic experiments have contributed much to the development of phonological theories, and they will continue to do so. In (3) and (4) I summarize how experiments have informed phonological theories and vice versa.

(3) How experiments inform theory

a. Provide data beyond those available through introspection.
b. Re-examine the quality of phonological data.
c. Address questions about the architecture of the grammar.
d. Show and examine the psychological reality of the grammar.
e. Find the sources of phonological patterns.

(4) How phonology informs experiments

a. Helps to find restrictions on – and the nature of – speech through phonological patterns.
b. Provides a guide as to where and what to look for in phonetic experiments.
c. Motivates specific hypothesis testing.

ACKNOWLEDGMENTS

Thanks to the audiences at Kyoto University, members at Kyushu University, especially Toshio Matsuura, and my students in my graduate seminar in Spring 2009 and the graduate phonetics class in Spring 2010 at Rutgers University, especially Aaron Braver, Will Bennett and Jimmy Bruno. Particular thanks are due to Osamu Fujimura, Beth Hume,
Michael Kenstowicz, Marc van Oostendorp, Jaye Padgett and two anonymous reviewers for their comments on earlier versions of this chapter, although they are not responsible for remaining errors. The preparation of this chapter is partially supported by a Research Council Grant from Rutgers University.

REFERENCES


