Kathleen Currie Hall

A typology of intermediate phonological relationships¹

Abstract: This paper presents an overview of phonological relationships that are “intermediate” between contrast and allophony. As has been observed for many years, such intermediate relationships occur widely in the world’s languages, yet they are often simply relegated to being exceptional cases or ignored in linguistic analyses. While there is a set of criteria that can be used to define relationships, these criteria may conflict with one another or be inadequate in certain cases. Thus, relationships may be intermediate for a number of reasons, including issues with: the concept of predictability of distribution; the existence of foreign or specialized strata of a language; variability and gradience; frequency; theory-internal classifications; and phonetic factors. Each of these is discussed in turn, with examples from the literature. In addition to this basic typology of intermediate phonological relationships, the paper also lays out a number of different kinds of approaches that have been taken to accommodate intermediate relationships in the phonological grammar and summarizes the issues to be resolved in future work.

Keywords: phonology, contrast, allophony, marginal phonemes

1 Introduction

1.1 The existence of intermediate phonological relationships

In phonology, one of the basic concepts is that not all differences in sounds are equally important. Some sounds – contrasting phonemes – can be used to distinguish among words, while others – allophones – are predictable variants of phonemes that mark positions within words. Thus there are classically two

¹ I am particularly grateful to Emmanuel Ferragne, Daniel Currie Hall, Beth Hume, and two anonymous reviewers for comments on earlier versions of this paper, as well as the audience of the 2012 Manchester Phonology Meeting. The usual caveats apply.
phonological relationships that can hold between pairs of sounds, *contrast* and *allophony*.

There are a number of criteria that have been used to determine the relationship between any pair of sounds, which will be described below in Section 2. There is not, however, an agreed-upon method for applying the criteria, and there are no guidelines for resolving cases in which the criteria conflict. Furthermore, there are in fact a large number of cases in which the criteria fail to satisfactorily classify relationships. This failure has led a substantial number of phonologists to refer, in both descriptive and theoretical work, to relationships that stand somewhere between contrast and allophony. I will refer to such relationships in this paper as *intermediate phonological relationships*. A sampling of the wide range of terms that have been developed to describe such situations is given below in (1); terms such as “marginally contrastive” or “marginally phonemic” are most common, but are certainly not the only ones.

(1) Terms in the literature for intermediate phonological relationships (relationships that are neither contrastive nor allophonic):
- *semi-phonemic* (e.g., Bloomfield 1939; Crowley 1998)
- *hemiphoneme* (e.g., Gleason 1961)
- *semi-allophonic* (e.g., Kristoffersen 2000; Moren 2004)
- *quasi-phonemic* (e.g., Scobbie et al. 1999; Hualde 2005; Vajda 2003; Gordeeva 2006; Scobbie and Stuart-Smith 2008)
- *quasi-contrastive* (e.g., Scobbie 2005; Ladd 2006)
- *quasi-allophonic* (e.g., Collins and Mees 1991; Rose and King 2007; Bye 2009)
- *quasi-complementary distribution* (e.g., Ladd 2006; Fougeron et al. 2007)
- *deep allophone* (e.g., Moulton 2003)
- *weak contrast* (e.g., Burzio 2001; Hume and Johnson 2001; Gormley 2003; Bennett 2005; Walker 2005; Wayment 2009; Giavazzi 2010; Martin and Peperkamp 2011; Shaw and Balusu 2010)
- *partial contrast* (e.g., Dixon 1970; Austin 1988; Hume and Johnson 2003; Frisch et al. 2004; Chitoran and Hualde 2007; Kager 2008)
- *semi-contrast* (e.g., Goldsmith 1995; Baković 2007)
- *just barely contrastive* (e.g., Goldsmith 1995)
- *fuzzy contrast* (e.g., Scobbie and Stuart-Smith 2008)
- *mushy phonemes* (e.g., Crowley 1998)
- *crazy contrast* (e.g., Boersma and Pater 2007)
- *effective contrast* (e.g., Ussishkin and Wedel 2009)
- *gradient phonemicity* (e.g., Ferragne et al. 2011; Boulenger et al. 2011)
There are many different reasons that a relationship may be intermediate between contrast and allophony. The purpose of this paper is to catalogue these various reasons and to outline the kinds of approaches that have been or may be taken to account for intermediate relationships. It should be noted that this paper does not propose a particular unique solution to this problem, in part because it may not in fact be possible (or in fact desirable; see Scobbie [1993: 52]) to have a single solution for all of the different cases of intermediacy. Rather, this paper is intended as a resource for researchers interested in the issue of intermediate phonological relationships – a guidebook, as it were, of some of the problems to be solved and some of the potential pathways to their resolution. Furthermore, this paper is designed to give a broad enough sampling to illustrate the proposed classifications and to demonstrate the wide-reaching nature of the problem of intermediacy across the world’s languages.

In addition to the list of those who have developed terms that have been used for intermediate relationships, Gleason (1961), Crothers (1978), Goldsmith (1995), Hill (1998), Hualde (2005), Ladd (2006), Scobbie and Stuart-Smith (2006), Kager (2008), Hall (2009, 2012), Bye (2009), Dresher (2011), Ferragne et al. (2011), and Boulenger et al. (2011) all explicitly discuss the existence of relationships that fall outside the domain of simple contrast and allophony. Gleason (1961: 298), for example, says that “[I]t will be necessary to do more than merely label phonologic
contrasts as ‘phonemic’ or ‘nonphonemic’, if we are to do all we should,” and Crothers (1978: 102) explains that “[c]riticisms [of the basic principle of phonemics] can be largely met, in [his] opinion, by adding to phonemic theory a distinction between marginal and full phonemes, the former including all phonemes with a severe distributional restriction in terms of phonological environment, or morphological or lexical function.” Hill (1998: 72) mentions in his discussion of some intermediate relationships in Texhuacan that “[m]arginal phonemes are important in dialectological work; it is a shame this important datum is dismissed so casually.” The other authors listed above are even more explicit in their observations and in some cases their proposals for dealing with such relationships. The current paper includes many of their observations and examples of intermediacy. However, where many of those papers focus on either a particular language (Hualde 2005; Scobbie and Stuart-Smith 2006) or a particular solution (Goldsmith 1995; Ladd 2006; Kager 2008; Hall 2009, 2012), the goal of the current paper is to compile a more comprehensive list of the ways in which phonological relationships may be intermediate and the ways in which intermediacy may be dealt with. That is, while there have been a handful of researchers who have explicitly recognized the issue of intermediate phonological relationships, and a wide range of researchers who have had to use some (often ill-defined) term or description of intermediate relationships when they come across them in a particular language, there has not yet been a concrete laying out of the problems that we face. This paper attempts to remedy this.

1.2 The problem of intermediate phonological relationships

Before delving into a discussion of intermediacy in phonological relationships, it is worth considering why the distinction between contrast and allophony is important to phonology and thus why relationships that cannot clearly be categorized as either contrast or allophony would in fact be problematic. In the days of structural analysis (e.g., Swadesh 1934, Twaddell 1957 [1935], Pike 1947), of course, a phonemic analysis of a language was an end in its own right, and many of the current criteria that are used for determining phonological relationships were developed in this period (see discussion in Section 2). For most modern phonologists, however, dividing a language’s sounds into those that are contrastive and those that are allophonic is simply a step that is taken along the way to understanding other aspects of a language’s sound patterns. At the same time, however, it is generally assumed to be a crucial step: as Ladd (2006: 10) points out, “[f]or a theoretical construct that was discredited forty years ago, the classical phoneme is actually still doing pretty well.”
With the advent of generative phonology (e.g., Chomsky and Halle [1968]; henceforth SPE), the focus of phonological analysis moved away from structural analysis and toward a theory of phonological rules in which there was no place for the notions of contrast and allophony as primitives. As Kenstowicz and Kisseberth (1979: 29) explain, however, the “null hypothesis” that the pronunciations of a set of words or morphemes is memorized and there is no need for generative phonology “fails to make a linguistically fundamental distinction between two types of phonetic information – namely, that some aspects of the phonetic realization are idiosyncratic properties of that particular morpheme, while others are instances of systematic regularities in the sound structure of the language” (emphasis original). This “linguistically fundamental distinction” is of course the difference between properties of segments that are contrastive (idiosyncratic, or unpredictable) and those that are allophonic (systematically regular, or predictable).

This difference was typically encoded through the use of phonological rules that applied to underlying representations to fill in all of the predictable information. Thus underlying representations were bundles of unpredictable features (representing contrastive information), and the rules encapsulated the predictable (allophonic) information. Subsequent years have seen the development of a number of different theories of underspecification that made various modifications to this basic distinction, including contrastive underspecification (in which all and only contrastive features are specified in lexical entries; e.g., Clements [1988]; Steriade [1987]), radical underspecification (in which any and all predictable information is left unspecified; e.g., Archangeli [1984, 1988]; Archangeli and Pulleyblank [1989]), and modified contrastive specification (in which the initial state is a single, undifferentiated phonological category (an undifferentiated “primordial allophonic soup” according to Dresher [2003b: 89]), and contrasts that are demonstrated to be present drive the specifying of particular features; e.g., Avery and Rice [1989]; Rice [1992]; Dresher [2003a, 2003b]; Hall [2007]; Mackenzie [2005]). In any of these theories of underspecification, however, there is a crucial difference between kinds of phonological information: that which is necessarily specified, and that which can be filled in by means of some other predictable generative process (e.g., rules). Interestingly, most of these theories assume that there are independent criteria for determining the difference between these two kinds of information (predictable / non-predictable, contrastive / non-contrastive, active / inactive, etc.). Regardless, some method of distinguishing among the various properties of phonological units is invoked, and any problems or indeterminacies associated with that method are important issues to be dealt with (see also discussion in Bye [2011] and an example of the problem in Moulton [2003]).
The difference between contrast and allophony is also important in Optimality Theory (OT; e.g., Prince and Smolensky [1993]). Unlike either pre-SPE phonology, which made use of phonological inventories to represent phonological relationships, or SPE-style phonology, which represents relationships through the use of underlying representations plus rules, OT represents relationships through constraint rankings. There is no phonological inventory per se, although it might be possible to reconstruct one on the basis of the winning output forms; the notion of an inventory, however, is generally irrelevant to OT. This is because OT is designed, through the principle of Richness of the Base, to always give a language-specific optimal output for a particular input form, even when that input contains non-native elements (as might be the case, for example, with a foreign borrowing).

Because OT does not have rules, it has no direct way of encoding predictable versus unpredictable information through the application of rules, as in SPE. Instead, this difference emerges through the relative ranking of faithfulness and markedness constraints. As Hayes (2004: 7) states, “in mainstream Optimality Theory, constraint ranking is the only way that knowledge of contrast is grammatically encoded: there is no such thing as a (theoretically primitive) ‘phoneme inventory’ or other restrictions on underlying forms.”

Interestingly, though there is no direct representation of relationships in OT, the use of constraint rankings to encode them actually results in more relationships than just contrast and allophony. Kager (1999) and Steriade (2007) both give factorial typologies of how constraints can interact to represent different phonological relationships in basic versions of OT. With three basic types of constraints – faithfulness (F), positional markedness (PM), and general markedness (M) – there are four distinct types of outputs. The first is full contrast: two segment types contrast in all environments; this is the result if the faithfulness constraint dominates both of the markedness constraints (schematized as $F \gg M, PM$; the ranking between the latter two constraints does not matter). The second is complete neutralization, or no contrast. This is the result of general markedness outranking both faithfulness and positional markedness; again, the ranking between the latter two constraints does not make a difference ($M \gg F, PM$). The third possibility is allophony; this results from ranking positional markedness over general markedness over faithfulness ($PM \gg M \gg F$). Finally, the fourth possibility is the positional neutralization of a contrast; this results from the ranking of positional markedness over faithfulness over general markedness ($PM \gg F \gg M$). These are the only four types of relationships that are predicted to exist with these types of constraints. Note that, unlike earlier models of phonology, OT treats neutralization of a contrast (either complete or partial) in the
same manner as both contrast and allophony; that is, as its own kind of phonological relationship.\(^2\)

All of the three basic theories of phonology outlined above – structural analysis, SPE-style rule-governed phonology, and OT – encode differences between different kinds of phonetic information. They do not all encode these differences in the same way or for the same reasons, or make the same predictions about the nature of such relationships or how many such relationships there might be. At the same time, some distinction along these lines is clearly an important aspect of doing phonological analysis.

Why is it an important aspect? As Ladd (2006: 10) comments, “[i]n practice – in speech therapy, in speech technology, in orthography design, in studies of language acquisition, and for that matter most descriptive work within Optimality Theory – the phoneme notion continues to be applied as if there were no serious problem with it.” In other domains of theoretical phonology, too, the distinction remains important: for example, the difference between contrastive and non-contrastive sounds or features is often thought to be crucial in understanding other phonological processes – only unpredictable / specified / contrastive sounds, for example, may be triggers or targets of other phonological rules (the contrastivist hypothesis; see, e.g., Hall [2007], Dresher [2008], and the discussion and references therein). Furthermore, the distinction can be used to understand how native speakers conceptualize their language. A number of studies have shown that there are perceptual differences between sounds that are contrastively related and those that are allophonically related (e.g., Jaeger 1980; Ohala 1982; Dupoux et al. 1997; Pegg and Werker 1997; Whalen et al. 1997; Fowler and Brown 2000; Harnsberger 2001; Peperkamp et al. 2003; Flagg et al. 2006; Kazanina et al. 2006; Pruitt et al. 2006; Boomershine et al. 2008; Johnson and Babel 2010), and perhaps even that there are differences in production (e.g., Cristià and Seidl, to appear). Furthermore, the difference has been claimed to be important in the process of both first and second language acquisition (e.g., for word segmentation; see, e.g., Lehiste [1960]; Umeda and Coker [1974]; Nakatani and Dukes [1977]; Church [1987]; Jusczyk et al. [1999]; Altenberg [2005]; Ito and Strange [2009];

\(^2\) One particularly interesting variant of OT for the present discussion is Flemming’s Dispersion Theory of Contrast (e.g., Flemming 2005; Steriade 2007). In this theory, OT constraints directly impose restrictions on pairs of sounds in contrast (through Minimum-Distance and Maximum-Contrast restrictions, grounded in perceptual factors), rather than indirectly deriving contrast and allophony through constraints on individual sounds. Flemming argues that distinctiveness constraints should replace faithfulness constraints because faithfulness constraints will never (by definition) motivate change, and yet there seem to be “enhancements” of contrasts. Note that in such a theory, there is a much more direct encoding of contrasts; it is thus even more important in this theory to know what the set of possible contrasts is in a given language.
Seidl et al. [2009]; Newman et al. [2011]). In order to properly develop any theory of phonology, then, it is important to understand the range of possible phonological relationships in the world’s languages in order to know what kinds of distinctions are necessarily encoded in phonological representations. Intermediate phonological relationships, as described in this paper, clearly provide a set of potentially problematic cases for phonological theories; it is thus important to understand the nature and scope of such intermediate relationships if one is to develop a comprehensive theory of phonology. I have presented them here for the most part in the guise of structural phonemic analysis, simply for ease of discussion and comparison across theories; it should be kept in mind, however, that the distinctions among such relationships are important for nearly all theories of phonology, regardless of how they are encoded.

The structure of the paper is as follows. In Section 2, I describe the criteria that have been developed to classify phonological relationships and point out some of the obvious problems and contradictions among them. In Section 3, I provide a typology of intermediate phonological relationships as described in the literature. In Section 4, I briefly sketch a number of different kinds of approaches that have been taken to accommodate intermediate relationships in the phonological grammar. Finally, in Section 5, I summarize the issues to be resolved.

2 Criteria for determining phonological relationships

2.1 The criteria

The most-cited criteria for determining the phonological relationship between two sounds, A and B, are listed below (see also discussion in Steriade [2007]; Dresher [2009: 72, 2011]; Hall [2011]). As a general rule, the first two criteria (predictability of distribution and lexical distinction) are considered the most important or primary criteria, while the others are secondary and often used in conjunction with the primary criteria in cases of conflict or uncertainty. In the descriptions below, I follow the traditional approach and assume that two sounds, A and B, must be either contrastive or allophonic in a language (i.e., if two segments are not contrastive, they are allophonic, and vice versa). For expository purposes only, I also assume in the following descriptions that each criterion is able to determine the relationship perfectly (in absence of other criteria). In actuality, it is certainly not true that any of the criteria can be used in all cases to define phonological relationships absolutely, as will be shown in detail below.
Predictability of distribution: Two sounds A and B are traditionally considered to be contrastive if, in at least one phonological environment in the language, it is impossible to predict which segment will occur. If in every phonological environment where at least one of the segments can occur, it is possible to predict which of the two segments will occur, then A and B are allophonic.

Example: Given the environment [b_t] (and other similar environments) in English, it is not possible to predict which of [i] or [u] will occur; both [bit] beat and [but] boot are real English words. Thus, [i] and [u] are contrastive in English. Given the environment [_eit] (and other similar environments), it is possible to predict that [l], and not [l], will occur in many varieties of English (e.g., Clark and Yallop 1995: 97), because [l] but not [l] occurs in syllable-initial position. Given the environment [tei_] (and other similar environments), it is possible to predict that [l], not [l], will occur, because [l] but not [l] occurs in syllable-final position. Thus, [l] and [l] are allophonic in these varieties of English.

Lexical distinction, or the commutation test: Two sounds A and B are contrastive when the substitution of A for B in a given phonological environment causes a change in the lexical identity of the words they appear in. If the use of A as opposed to B causes no change in the identity of the lexical item, A and B are allophonic.

Example: Given the word beat [bit], substituting [u] for [i] changes the lexical identity to boot, [but]. Based on this criterion, [i] and [u] are contrastive in English. Given the word late [leit], substituting [l] for [l] does not change the lexical identity of the word (though the pronunciation might be considered slightly odd). Similarly, given the word tale [tei], substituting [l] for [l] does not change the lexical identity of the word. According to this criterion, then, [l] and [l] are not contrastive and are therefore allophonic in English.

Native speaker judgment: Two sounds A and B are contrastive if native speakers think of them as “different” sounds; they are allophonic if native speakers think of them as the “same” sound (or variations on the same sound). Such judgments may be tested directly, at the conscious level, or indirectly and subconsciously, through psycholinguistic experimentation.

Example: Native speakers of English readily identify [tʰ] and [pʰ] as distinct sounds in English; therefore, [tʰ] and [pʰ] are contrastive. Native speakers are usually unaware that there are two different versions of [t] ([t] and [tʰ]); hence [t] and [tʰ] are allophonic.
Example: Native speakers of English are faster to discriminate [d] and [ð] than they are to discriminate [d] and [ɾ], whereas native Spanish speakers show the reverse (Boomershine et al., 2008). Therefore, [d] and [ð] are contrastive in English and allophonic in Spanish, while [d] and [ɾ] are contrastive in Spanish and allophonic in English.

(5) Alternations: Two sounds A and B are contrastive if they participate in morphophonemic alternations with each other. A and B are allophonic if they participate in allophonic alternations with each other.3

Example: The plural morpheme /z/ in English is realized as [s] after voiceless non-sibilants (e.g., cats [kæts]), but as [z] after voiced non-sibilants (e.g., dogs [dɒgz]). This alternation neutralizes the phonemic difference between [s] and [z]; therefore, [s] and [z] are contrastive in English. The morpheme write /rait/ is realized with a [t] when it occurs in isolation (e.g., write [ræt]), but with a [ɾ] when it occurs as the first syllable of a trochaic foot (e.g., writer [raɪtər]). This alternation between [t] and [ɾ], which are not otherwise thought of as phonemic, indicates that they are allophonic in English.

(6) Phonetic similarity: Two sounds A and B cannot be considered allophonic if they are not somewhat phonetically similar; if they are not phonetically similar, they are considered contrastive.

Example: [tʰ] and [t] are predictably distributed in English ([tʰ] occurs syllable-initially and [t] occurs after [s]). They are phonetically similar according to subjective observation (e.g., they are both pronounced with an alveolar place of articulation); thus, they can be considered allophonic. [pʰ] and [t] are predictably distributed in English ([pʰ] occurs syllable-initially and [t] occurs after [s]). They are not phonetically similar according to subjective observation (e.g., one is bilabial and one is alveolar) and therefore cannot be considered allophonic; they must instead be considered contrastive.4

3 This criterion is obviously circular, as stated here. There is generally no clear way of distinguishing morphophonemic from allophonic alternations, except by means of the other criteria for determining contrast. See further discussion of its usual application in Section 2.2.

4 Note that there is no a priori reason to assume that place of articulation is a more important criterion for determining phonetic similarity than, for example, manner, voicing, or aspiration. Arguments based on phonetic similarity are almost always highly subjective in nature. See discussion in Section 2.2.
(7) **Orthography:** In a language with a phonographic writing system, two sounds A and B that are typically written with distinct graphemes are contrastive. Two sounds that are typically written with the same grapheme are allophonic.\(^5\)

- Example: In English, the segments [\(\text{th}\)] and [\(\text{ph}\)] are typically written with the distinct graphemes ⟨t⟩ and ⟨p⟩. Thus, [\(\text{t}\)] and [\(\text{p}\)] are contrastive. There is only one grapheme, ⟨t⟩, that is used to represent both [\(\text{t}\)] and [\(\text{t}\)]; thus, [\(\text{t}\)] and [\(\text{t}\)] are allophonic.

(8) **Place in the system / Phonological behavior:** If a sound A patterns in a way that is similar to other elements in the system that are believed to be contrastive, then A should also be considered contrastive in the system. (Note that depending on the level of abstractness allowed, A may be considered a contrastive phoneme in the language even if it is never realized on the surface.) If a sound A can be shown to be derived by means of some other regular pattern in a language, then it should not be considered a separate contrastive phoneme but rather an allophone of some other contrasting phoneme.

- Example (from Hyman [1970]): In Nupe, there is an apparent surface contrast among palatalized, labialized, and plain consonants before /a/ (e.g., [\(\text{tyá}\)] ‘to be mild’, [\(\text{twá}\)] ‘to trim’, and [\(\text{tá}\)] ‘to tell’). Elsewhere, palatalized and labialized consonants appear as predictable (allophonic) variants of plain consonants before non-low front and non-low rounded vowels, respectively. Therefore, one can posit /ɛ/ and /ɔ/ as contrastive phonemes in Nupe, which similarly trigger palatalization and labialization, but which are neutralized on the surface to [a], so that the underlying forms of the above words are /tɛ/ ‘to be mild,’ /tɔ/ ‘to trim,’ and /ta/ ‘to tell.’

- Example: In English, there is an apparent surface contrast between [\(\text{n}\)] and [\(\text{ŋ}\)], as in sin [\(\text{sin}\)] vs. sing [\(\text{sin}\)]. Coronal nasals are generally known to assimilate to the place of articulation of the following consonant, however, and many surface instances of [\(\text{n}\)] occur before velars or in words where there was historically a velar segment following them (sometimes this history is even visible through synchronic alternation, as in long [\(\text{n}\)] ~ longer [\(\text{ŋ}g\)]). Therefore, [\(\text{n}\)] can be assumed to be derived

---

5 Note that this criterion is related to criterion (4) about native speaker judgment, in that judgments of native speakers of languages with phonographic writing systems are quite likely to be influenced by those writing systems.
from an underlying sequence of /ng/ or /nk/ and should not be assigned status as a contrastive phoneme in English. It is instead an allophone of /n/.

Most of these criteria were developed in the first half of the twentieth century as part of the endeavor of phonemic analysis. The goal of phonemic analysis was to divide a language’s sounds into phonological categories – phonemes – so as to have an inventory of the categories available for symbolic manipulation in the language. Underlying phonemic analysis is the so-called phonemic principle, defined by Swadesh (1934: 118) as the facts that “there are in each language a limited number of elemental types of speech sounds, called phonemes, peculiar to that language; that all sounds produced in the employment of the given language are referable to its set of phonemes; [and] that only its own phonemes are at all significant in the given language.”

There were, however, many different views of the essential nature of the phoneme; for example, some thought it was a psychological entity (e.g., Sapir 1925, Swadesh 1934), others a physical one (e.g., Bloomfield 1933, 1935), and others an abstract organizational construct (e.g., Twaddell 1957 [1935], Trubetzkoy 1969 [1939], Pike 1947). Along with these different views of the nature of the phoneme were a number of different hypotheses about the proper procedures for conducting phonemic analysis, as pointed out by Chao (1957 [1934]) (e.g., Swadesh 1934, Twaddell 1957 [1935], Pike 1947). Chao (1957 [1934]) notes that these different procedures may result in different understandings of the phonological system, each of which may be useful under different circumstances. For example, Pike’s (1947) approach is as follows: “in order to be considered submembers of a single phoneme, two segments must be (a) phonetically similar and (b) mutually exclusive as to the environments which they occur” (62). He adds that in order for two segments to be considered members of different phonemes, they must “consistently constitute the only difference between two words of different meanings” (62). This approach clearly incorporates the criteria of predictability of distribution, lexical distinction, and phonetic similarity, but makes no reference to native speaker judgments, as for Pike, phonemes were simply abstract organizing units and not psychological realities. Swadesh’s (1934) approach is quite similar, but he also includes the following “test of substitution” to determine whether two phonetically distinct sounds belong to separate phonological categories: “... [P]ronounc[e] a word with some modification in one of the phonemes. If the modification cannot be perceived by a native, it is within the range of normal deviation. If the modification seems to trouble the native, it is an extreme deviation from the norm, a distortion. If the native definitely hears some other word or feels that one has the word wrong, one may conclude
that the modification has amounted to the substitution of one phoneme for another” (124). This is of course the test for native speaker judgment, which was important to Swadesh because of his belief that phonemes had psychological reality. Such conscious distinctions are rarely, however, considered today (though as noted above, experimental tests for the implicit psychological reality of phonemes have continued to play an important role in phonological research). Indeed, today the criteria tend to be blended together without regard to the particular nature of the phoneme that each was designed to identify, and there is very little said about what to do in instances in which the criteria conflict with one another.

### 2.2 Problems with these criteria

As is evident even from the descriptions of these criteria, situations can arise in which the criteria fail to produce easily interpretable results, either because a given criterion is insufficient or because multiple criteria conflict with each other. For example, the criterion of orthography is inadequate in languages without a one-to-one mapping between sounds and graphemes (though it admittedly is not called upon frequently; see Twadell [1957 [1938]] and Swadesh [1941] for examples of its use, and see Fischer-Jørgensen [1975: 4] for further discussion). Similarly, the criterion of alternation is insufficient because it is circular; by itself, it cannot be used to determine whether two sounds are allophonic or contrastive, because the types of alternations are simply defined in terms of allophony and contrast. Silverman (2006: 88) does make the claim that “the only way sounds can be allophonically related is if they alternate with each other, that is, if they partake in a sound substitution which maintains meaning” (emphasis original); Lu (2012) presents experimental evidence that tends to support this claim.

---

6 It should be noted that if there is an independent means of identifying rules that apply to allophonic vs. contrastive sounds, then it is possible to use the criterion of alternations to differentiate the two. In Lexical Phonology (e.g., Kiparsky 1982, 1993) for example, contrastive pairs of sounds tend to be related to each other through cyclic lexical rules that can make reference to derived environments (through interaction with various levels of morphology), while allophonic pairs are related through non-cyclic, post-lexical, rules that refer only to phonological environments. In this theory, then, alternations that are governed purely on the basis of differences in the phonological environment could be labelled as allophonic, while alternations that refer to morphological boundaries or other derived environments could be labelled as morphophonemic (contrastive).
that this is phrased as a requirement of allophony, not as a diagnostic: for sounds to be allophonic, they must alternate, but not all alternating sounds are claimed to be allophonic.

The criterion of lexical distinction, one of the apparently most robust criteria, is insufficient in that it seems to be stronger at identifying contrastive than allophonic relationships: that is, it is more clearly the case that the existence of a minimal pair should be an argument for contrast than that the lack of a minimal pair should be an argument for allophony, given the possibility of “accidental gaps” (and indeed, phonologists commonly accept “near” minimal pairs and “possible” words of a language as evidence for contrast in the absence of actual minimal pairs). What is more, there are cases when the substitution of a particular sound that is generally considered an allophone of a particular phoneme may in fact change the percept of the phoneme category, contra to the prediction of lexical distinction. This is the case with $[pʰ]$ and $[p]$, which contrast in English in word-initial position for speakers who pronounce bit as $[pɪt]$, but are non-contrastive after $[s]$ – both $[spʰɪt]$ and $[spɪt]$ would be accepted as the word spit.

Furthermore, criteria can conflict with each other, as in the case of predictability of distribution and phonological patterning. Based on predictability of distribution, for example, there is an apparent surface allophony in English between $[h]$, which predictably appears syllable-initially, and $[n]$, which predictably appears surface-finally. As was shown earlier, however, based on other patterns of nasal place assimilation, it might be claimed that $[n]$ is derived from underlying sequences of $/n/ plus a velar. Then, $[h]$ and $[n]$ could not be allophones of one another because both $/h/ and /n/ (of which $[n]$ is simply a derived allophone) can occur syllable-initially.

A similar conflict can be seen between predictability of distribution and phonetic similarity. Generally speaking, the criterion of phonetic similarity is used as a filter on the criteria of predictability of distribution and lexical distinction. Phonetic dissimilarity has also been called upon to explain why $[h]$ and $[n]$ are not allophonic in English, or to explain why $[t]/[tʰ]$ are allophonic while $[t]/[pʰ]$, which are also predictably distributed in English, are not. Janda [1999] takes a particularly strong view on the role of dissimilarity, claiming that a set of sounds that is in complementary distribution and analyzed as allophonic can undergo phonemic split if they simply undergo sound changes that make them sufficiently phonetically dissimilar. Mannell (2008), however, takes the opposite strong view that “The[re] are no obvious and consistent acoustic, auditory or articulatory criteria for phonetic similarity,” and Silverman (2006: 87) claims that “articulatory or acoustic similarity between sounds is neither a prerequisite, nor a diagnostic, for allophonic relatedness” (emphasis original). Perhaps this is because no such consistent criteria are possible, but the lack of clear guidelines for when and how
phonetic similarity can or should be used in conjunction with other criteria – even if its use is necessarily variable – can lead to confusion and discrepancies in the analysis of phonological relationships.

In addition to these relatively straightforward problems with the criteria, there are also a number of more subtle difficulties in applying the criteria in particular cases (e.g., knowing what counts as “predictable”). Furthermore, there appear to be other factors that are not included in these criteria that phonologists believe to be relevant to the actual determination of phonological relationships (e.g., frequency of occurrence). These difficulties are the ones that seem to have given rise to the wide range of terms for intermediate relationships given in Section 1. In the following section, I provide a typology of these problems, relating them back to the criteria listed here.

3 A typology of intermediate phonological relationships

This section lays out a typology of intermediate phonological relationships, based on the observations and descriptions made in the literature about relationships that are simply not clearly classified by the criteria given in Section 2.1. The reasons for intermediacy are varied, but can broadly be classified under the following headings:

– problems with the concept of “predictability of distribution” (Section 3.1)
– foreign, specialized, or otherwise distinct strata of languages (Section 3.2)
– issues of variability and gradience (Section 3.3)
– issues of frequency (Section 3.4)
– subsets of natural classes (Section 3.5)
– issues of phonetic similarity (Section 3.6)

Finally, Section 3.7 discusses a few cases of intermediate relationships that are highly complex and do not fit neatly into this typology. In general, it should be noted that the meaning of a term such as “marginal” or “quasi-” is not always made explicit in descriptions that use them and that multiple meanings are sometimes collapsed in the same discussion; the typology that follows is based on close readings and my own interpretation of the authors’ observations.
3.1 Problems with the concept of “predictability of distribution”

The basic concept of predictability of distribution is quite straightforward; if the choice between two sounds can be predicted from their phonological environment, they are allophonic, but if there is at least one instance in which the choice is unpredictable, then the sounds are contrastive. When it comes to applying this criterion, however, there are numerous problems that arise: a pair of sounds may be mostly unpredictable, but have some degree of predictability in their distribution, or may be mostly predictable, but with a few contrasting environments. While straight application of the criterion would dictate that such cases simply be considered contrastive because they do not involve complete predictability, phonologists have often shied away from this analysis because of intuitions that it obscures relevant facts about the language. Another problem with the criterion is that it is not always clear what “counts” in terms of defining predictability; for example, should morphological boundaries count in defining phonological environments? Each of these problems will be considered in turn, with examples of intermediate phonological relationships that arise because of them.

3.1.1 Mostly unpredictable, but with some degree of predictability

Perhaps the most well-known type of intermediate relationship is the case where two phonological units (segments, features, prosodic structures, etc.) are contrastive in most cases, but are predictable in one or two others – these are in fact cases of phonological neutralization.

As a general proposition, a contrast that is neutralized in a particular environment is still considered contrastive and thus not particularly problematic. Trubetzkoy (1969 [1939]: 239) points out, however, that neutralization can lead to either a slight or a severe reduction of the “distinctive force” of an opposition. Some researchers have indeed interpreted neutralization as creating a separate relationship somewhere between full contrast and full allophony. Hume and Johnson (2003), for example, refer to neutralized contrasts as “partial contrasts” and give experimental evidence supporting Trubetzkoy’s hypothesis, showing that Mandarin tones that are partially contrastive are perceived as being more similar than fully contrastive tones. Similarly, Kager (2008) also refers to contextual neutralization with the term “partial contrast,” again suggesting that contrastive relationships are not all created equal and that neutralization of a contrast changes the relationship in some fundamental way (and indeed, as described
above in Section 1.2, Optimality Theory in general accounts for neutralization and allophony using the same mechanisms.

To take another specific example, Hualde (2005) describes the classic problem of the distribution of the trill [r] and the flap [ɾ] in Spanish as an example of a “quasi-phonemic” relationship. He concludes that the two segments are contrastive because of the robust presence of minimal pairs where [r] and [ɾ] contrast intervocally, and that the contrast is simply neutralized everywhere else. He adds the caveat, though, that “[r] and [ɾ] are clearly more ‘closely related’ than other pairs of phonemes” (Hualde 2005: 19–20).

Ladd (2006) reports a similar type of “close” relationship resulting from the neutralization of higher and lower mid vowels in French and Italian. The vowels contrast only in lexically stressed syllables and are neutralized elsewhere; Ladd refers to this as being a “quasi-contrastive” relationship.

One classic way of handling neutralization in a formal framework is with the archiphoneme (e.g., Trubetzkoy 1969 [1939]), a unit whose distinctive features are those that are shared by two neutralized phonemes and that distinguish those phonemes from the rest of the inventory of the language. For example, if otherwise contrastive /p/ and /b/ are neutralized in word-final position, one would claim that neither /p/ nor /b/ occurs in word-final position but rather the archiphoneme /P/, which would simply be an oral bilabial stop with no voicing specification. As Anderson (1985) discusses, Trubetzkoy presents archiphonemes as a separate unit within the phonological system, implying some special status for neutralized elements, but that representation has not persisted, with the archiphoneme being more simply a shorthand for the common set of features that occur in a neutralized position. At the same time, the fact that some pairs of phonemes are linked through an archiphoneme in some languages while other pairs are not, or are not in other languages, leaves open the possibility that these pairs are different in some sense, as they do indeed have a distinct formal representation.

Thus, despite the common assumption that contrastive pairs are still contrastive even when they can be neutralized, there is at least some feeling that the neutralization actually changes the nature of the phonological relationship in some fundamental way; that is, that neutralized contrasts are distinct from both contrastive and allophonic relationships.

### 3.1.2 Mostly predictable, but with a few contrasts

The same problem of “incompleteness” can be found with relationships that are basically allophonic, but seem to be unpredictable in a few environments. Of particular note is that while there is nothing inherently different about the end
result of such examples from that of the examples given just above in Section 3.1.1
(both are cases where pairs of sounds are predictable in some environments
and unpredictable in others), there is a tradition of distinguishing between relationships that are basically contrastive, but neutralized (Section 3.1.1) and those
that are basically allophonic, but with a few contrasts (this section). It is often the
case that the distinction between the two is actually diachronic; a synchronic interpretation of “neutralized contrast” is given when there was historically a
contrast in a language, while a synchronic interpretation of “basically allophonic” is given when there was historically a completely predictable relationship. Goldsmith (1995: 10) distinguishes these two and places them at different points on a “cline” of contrast – those where the basic pattern is contrastive
but neutralized are cases of “modest asymmetry” or “not-yet-integrated-semi-
contrasts,” while those where the basic pattern is one of predictability are cases
of being “just barely contrastive” (see also Section 4.4).

Within the category of basically predictable relationships that show some
degree of contrastiveness, there are in fact two sub-categories: those where the
few contrasts are systematic and those where they are exceptional (i.e., lexical
irregularities). Examples of systematic unpredictability are particularly difficult
to distinguish from neutralized contrasts. Hualde’s (2005) “quasi-phonemic” ex-
ample of Spanish [ɾ] and [r] discussed above exemplifies this point. Although
Hualde concludes that the two are contrastive, because of intervocalic contrasts,
he points out that the two are in fact predictably distributed elsewhere. One could
thus imagine that a different analyst might have come to a different conclusion:
that the widespread predictability makes these two sounds allophonic, but with
a few exceptional contrastive environments.

Another example of systematic unpredictability is that of Canadian Raising,
a phenomenon that has been reported for many dialects of English, both within
and outwith Canada (e.g., Joos 1942; Chambers 1973, 1975, 1989; Trudgill 1985;
Vance 1987a; Allen 1989; Britain 1997; Trentman 2004; Fruehwald 2007). The
diphthongs [ɑɪ] and [ɑ][, or [ɑʊ] and [ʌ],[ are generally predictably distributed,
with [ɑ] and [ʌ] occurring before tautosyllabic, tautomorphemic voiceless seg-
ments, and [ɑ] and [ʌ] occurring elsewhere (e.g., tight [tʌɪt] but tide [tɑɪd];
lout [lʌut] but loud [lɑʊd]). There are, however, surface (near) minimal pairs dis-
tinguished by the two vowels, such as writing [rɑɪɾɪŋ] and riding [rɑɪɾɪŋ], or outer
[ʌʊɾɚ] and louder [lɑʊɾɚ], in which the two systematically contrast before a
flap [ɾ]. Given the presence of such minimal pairs, it has been argued that [ɑ] and
[ɑ], and [ʌ] and [ʌ], are contrastive in Canadian English (and other similar dia-
lects) (see, e.g., Kiparsky and Menn 1977; Mielke et al. 2003), but others have been
reluctant to relinquish the status of the two as allophonic, largely because the
pattern is actively productive in nonsense words (e.g., Bermúdez-Otero 2003;
Boersma and Pater (2007; Idsardi 2006) (or because the process of flapping itself is assumed to be predictable, thus allowing the vowel quality to be predicted from the underlying representation). Myers (1997: 173), pointing out further minimal pairs such as spider [ʌɪ] vs. cider [ɑɪ], with no obvious morphological basis for the distinction, characterizes this situation as “marginally contrastive.” We will return to this example in Section 3.1.3; see also Bye (2011).

Other examples of systematic exceptions to basically predictable relationships can be found in, for example, Gothic ([i]/[j] and [u]/[w]; Kiparsky [2003: 6]), Gugu-Badun and Biri (lamino-dentals vs. lamino-palatalis; Dixon [1970: 93]), Korean (vowel length; Kochetov [2008a: 161]), Menomini (long [ū]/[ō]; Bloomfield [1939: Section 35]), and Rejang ([a]/[a]; Blust [1984: 426]).

In addition to these systematic deviations from total predictability of distribution, there are many cases where the deviation is irregular – for example, caused by lexical exceptions. Examples of this include the classic cases of /æ/-tensing in New York City and Philadelphia (e.g., Labov 1981, 1994), in which, for example, lax /æ/ occurs before voiced stops (e.g., in cad [kæd]) while tense /æ/ occurs before voiceless stops (e.g., in cat [kæt]). This is true except in a particular set of words including mad, bad, and glad, in which a tense /æ/ always occurs ([mæd], [bæd], [glæd]; see Labov [1994: 431]). Moren (2004) describes this pattern as being “semi-allophonic,” and Bauer (1995) says that this contrast “is marginal in the English phonemic system” because of the facts that (1) speakers are generally unaware of the difference and cannot use it consciously to distinguish words and (2) the distribution is mostly predictable.

The case of long [ū] in Menomini, mentioned above, also contains lexical exceptions. In borrowed words, [o̞] and [ū], which are normally predictably distributed, can contrast as in [cōh] ‘Joe’ versus [cūh] ‘Jew’ (Bloomfield 1962: Section 1.16). Other examples of lexical exceptions to otherwise allophonic patterns can be found in Chaha ([n]/[r]; Banksira [2000], Rose and King [2007]), Denjongka of Sikkim (vowel length; Yliniemi [2005: 45]), Enets (vowel length; Anderson [2004: 25]), English (full vowels vs. schwa; Chomsky and Halle [1968]; Ladd [2006: 14]), Frenchville, PA, French ([ø]/[œ]; Bullock and Gerfen [2005: 120]), Jicaque ([β]/[w]; Campbell and Oltrogge [1980: 211]), Korean ([i]/[ni]; Sohn 2008), Modern Greek (voiced stops; Viechniki 1996), Polish ([j]/[g]; Zygis and Padgett [2010: 209, 210]), and Spanish ([j]/[i]; Hualde [2005]).

As a last example in this section, consider Janda’s (1999) argument that the presence of even complete predictability of distribution does not entail allophony. His focus is on phonological change; Janda (1999: 330) points out that phonemic splits often give rise to what he refers to as “marginal/quasi-/secondary phonemes.” His use of the term refers to segments that are descriptively in complementary distribution (and hence would normally be classified as
being allophonic) but must be considered by native speakers to be separate phonemes, as evidenced by later loss of the conditioning environments but preservation of the distributions. Janda gives as an example Twaddell’s (1957 [1938]) (and others’) account of the historic change of umlaut in German. According to Janda’s description, in Old High German, the back rounded vowels /o/ and /u/ were predictably realized as front rounded vowels when they were followed by a front vowel in the next syllable; in other words, the back and front rounded vowels were in complementary distribution and entirely predictable. At some point, however, front vowels in final syllables were lost, so the triggering environment for fronting of rounded vowels was lost. Front rounded vowels remained in the words where they had originally been conditioned, however. The traditional conclusion is that the distinction between front and back rounded vowels must have been phonemicized even while the predictable environments were still there – that is, while they were still in complementary distribution (assuming, of course, that the loss of the front triggering vowels did not happen precisely simultaneously with the contrastive reanalysis of front and back rounded vowels). Thus, even sounds that are predictably distributed may be analyzed as being “marginal” or “secondary” phonemes if there is reason to believe that speakers have treated them that way.

3.1.3 Defining predictability

In addition to these instances where it simply isn’t clear that predictability of distribution should be an all-or-nothing criterion, it is often unclear exactly how to define what makes a pair of sounds predictably distributed. There are a number of cases in which segments in a language are in fact predictably distributed, but this predictability is evident only when non-phonological factors (e.g., morphological or syntactic) are considered. This means, then, that the criteria of lexical distinction and predictability of distribution can conflict with each other, despite the fact that they are often quite similar and both give rise to the minimal pair test for determining contrasts.

As an example of how these two can conflict if predictability of distribution is not clearly defined, Scobbie (2002) describes cases where two segments are the only audible contrast between words, but are predictably distributed if morphological information is taken into account. For example, the distinction between [ai] and [ʌi] in Scottish English is the only audible difference between the words tied [tæd] and tide [tʌd]; the minimal pair test as based on lexical identity should dictate that the sounds [ai] and [ʌi] are contrastive. However, when the morphological boundaries of the two words are considered, the use
of [ɑɪ] as opposed to [ʌɪ] is predictable: [ɑɪ] is used morpheme-finally (tie + d) while [ʌɪ] is used morpheme-internally before a stop (tide). The same pattern holds true of the entire distribution of these two vowels; under the criterion of predictability of distribution, then, these two segments should be considered allophonic. Harris (1994) gives examples of similar cases, such as the difference between pause [powz] and paws [pɔəz] in London English, between daze [dɪəz] and days [dɛːz] in northern Irish English, and between the vowels of ladder [lædə] and madder [meədə] in New York and Belfast English (see also Bauer [1995]). The question is whether morphological information should be allowed to “count” toward determining the predictability of the distributions, a question that is left unanswered by the criteria above and which has received different answers over the years. Such cases are also given the names “quasi-contrast” (e.g., Ladd 2006) or “fuzzy contrast” (e.g., Scobbie and Stuart-Smith 2008); note that Carr’s (2008) Glossary of Phonology in fact defines the term “marginal contrast” as a “derived contrast” with specific reference to this kind of morphologically complex phenomenon.

Bye (2009) refers to cases like these as instances of “morphological override,” indicating that although a purely phonological approach would argue for one analysis on the basis of predictability of distribution, morphological factors can supersede. His example is that of Javanese, where “the distribution of ATR and RTR high vowels appears largely governed by syllable structure (ATR in open syllables, RTR in closed syllables) . . . . [but in the] so-called elative (intensive) form of the adjective, however, [there is] ablaut of the final vowel to a high ATR vowel regardless [of] whether the syllable is open or clased” (Bye 2009: abstract). Note that here, the same word may have multiple pronunciations, depending on the particular morphological form; this is somewhat different from the above cases where it is simply the presence or absence of a morphological boundary that matters.

Another classic example of a non-phonologically conditioned contrast is the [x] ~ [ç] distinction in German. These two voiceless fricatives are generally predictably distributed, with [x] appearing after a low or back vowel (e.g., ach [ax] ‘oh’) and [ç] appearing elsewhere (e.g., ich [ic] ‘I’). Despite this pattern of predictability, [x] and [ç] in German constitute one of the best-known cases of a “marginal” contrast, as there are both native German words and borrowed words where the usual distribution of the pair does not hold (e.g., Kuchen [ku.xən] ‘cake’ vs. morphologically complex Kuhchen [ku.çən] ‘little cow’; see Fox [1990] for other examples). It has long been debated whether minimal pairs such as these are sufficient to establish [x] and [ç] as being contrastive in German; Robinson (2001) gives a comprehensive description of the arguments and analyses on either side. Fox (1990: 41) sums up the reasoning of proponents of the allophonic
analysis, saying that “it seems undesirable – and, one might add, against the feeling of the native German speaker – to complicate our analysis [by establishing /ç/ as a separate phoneme], especially as the relationship between these two sounds is otherwise such a clear case of complementary distribution.” It is clear from this debate that applying the criterion of predictability of distribution is not straightforward and that issues about morphological conditioning and native speaker intuition (a separate criterion for determining phonological relationships in the first place) also come into play.

An additional problem with the definition of predictability of distribution is described in Dresher (2011: 253) and referred to as “partial phonemic overlapping.” In overlapping cases, two phones can occur in the same phonological environment, and thus would appear to be unpredictably distributed (contrastive). At the same time, the same two phones can also be shown to be predictably distributed in certain environments and thus appear to be allophonic. An example is Chukchee, in which the two sounds /i/ and /e/ can contrast with one another (e.g., [meməl] ‘seal’ vs. [miməl] ‘water’; see Spencer [1999]), and yet [e] also appears as a predictable allophone of /i/ when a low vowel appears elsewhere in the word (e.g., [maməl-ma] ‘with seal’ vs. [meməl-ma] ‘with water’). Thus, [e] appears to be an allophone of two separate phonemes, /i/ and /e/; it is predictably an allophone of /i/ if there is a low vowel elsewhere in the word and predictably an allophone of /e/ if there is no such low vowel. Simply looking at their surface distributions, however, one would conclude that [e] is unpredictably distributed with respect to [i] and therefore that the two are contrastive; this does not, however, elucidate the whole story. Such examples highlight the fact that predictability of distribution is not as straightforwardly defined as one might think and in particular that elements and relationships may be “predictable” on different levels.

A related problem with the application of the criterion of predictability of distribution is that sounds may have different distributions at different levels of analysis, and there is not a consensus about which stage should be used when applying the criteria to make decisions about phonological relationships. In many theories of phonology, it is assumed that phonological operations act to map an underlying representation onto a surface representation (with varying levels of intermediate representations allowed). As mentioned in Section 3.1.2, the distribution of [ɑɪ] and [ʌɪ] in Canadian English is therefore problematic.

7 The forms containing [ma] are hypothesized for illustrative purposes from descriptions of Chukchee morphophonology in Kenstowicz & Kisseberth (1979), Spencer (1999), and Dresher (2011), but have not been verified by native speakers.
On the surface – that is, in spoken language – the distribution of [ɑɪ] and [ʌɪ] in Canadian English is unpredictable in at least one phonological environment, namely, before [ɾ], resulting in minimal pairs like rider [ɹɑɪɾɚ] and writer [ɹʌɪɾɚ].

Thus, on the basis of the criteria of predictability and lexical distinction, these two sounds should be considered contrastive. If one assumes, however, that [ɾ] is not present in the underlying representation and is simply a derived allophone of both /t/ and /d/, the distribution of [ɑɪ] and [ʌɪ] becomes predictable at the underlying level of representation. [ʌɪ] occurs before tautosyllabic voiceless segments, while [ɑɪ] occurs elsewhere. If this is the case, then the two diphthongs should be considered allophonic. The choice of using surface representations or underlying representations in determining distribution, then, has consequences for the ways in which sounds are assigned to phonological relationships, but there is no criterion to determine which level of representation to use. It may, of course, simply be the case that there is no unique right answer, and that the answer just depends on why one is identifying the phonological relationships in the first place (cf. Section 1.2). The outcomes of such issues are, however, commonly debated in the phonological literature, and it would be useful to resolve these debates either by coming up with a unique answer or by more explicitly identifying the instances in which one answer is better than another.

### 3.2 Foreign, specialized, or otherwise distinct strata of languages

The preceding section has described a number of intermediate phonological relationships that can arise from different interpretations or applications of the criterion of predictability of distribution. There are, however, a number of other reasons for which phonological relationships may be considered intermediate. The first of these is that there are often multiple linguistic strata in a language (a situation sometimes referred to as polysystemicity), and the criteria may give different results when applied to different strata. These strata may arise from

---

8 It should be noted that there are further complications to this distribution, in the form of high and low variants appearing in contexts not predicted by phonological rule (e.g., Hall 2005). Even without these additional complications, however, Canadian Raising poses problems for traditional definitions of contrast and allophony.

9 Note that this does not explain the occurrence of minimal pairs such as spider [ʌɪ] vs. cider [ɑɪ], as pointed out by Myers (1997).
historical influences on the language (e.g., borrowings from other languages
that include a borrowing of some of that language’s phonological principles)
or from other differentiation within the language (e.g., the development of
specialized language usage for religious or other purposes).

For example, Ladd (2006) reports that there are several indigenous Mexican
languages where voiced stops are usually allophonic, but which are beginning to
have contrastive voiced stops through contact with and borrowing from Spanish.
As a more specific example, in English, [s] and [ʃ] are considered to be contrastive
on the basis of minimal pairs like sue [su] and shoe [ʃu], mass [maes] and mash
[mæʃ], etc., which indicate contrastivity by the criteria of lexical distinction,
predictability, and orthography. In initial consonant clusters, however, their distribu-
tion is largely predictable: [s] appears before [i] while [s] never does (e.g.,
shriek [ʃiik], *[siiik]); [s] appears before other consonants, while [ʃ] never does
(e.g., sleep [slip], *[ʃlip]; school [skul], *[ʃkul]). While this might be taken as an
example of contrast neutralization, the situation is complicated by the existence
of borrowed words from Yiddish with [ʃ]-consonant clusters; for example, schlep
[ʃlep], schmooze [ʃmuz], spiel [ʃpiɫ]. These are all “foreign” words at some level,
with native English speakers varying in their knowledge of and acceptance of the
words. These borrowings have even resulted in a minimal pair: stick [stɪk] vs.
schtick [ʃtɪk]. The question, then, is whether [s] and [ʃ] are simply entirely con-
trastive (there being minimal pairs in all positions in at least some stratum of
the language) or a contrast that is subject to neutralization, because in some
native strata there is occasional predictability. Again, it should be empha-
sized that there may not be a single “right” answer, but this example highlights
some of the issues with trying to categorize relationships as either contrastive
or allophonic.

Perhaps the most well-known example of this phenomenon is in Japanese,
which is traditionally assumed to have four different morpheme classes that have
their own phonological patterns: Yamato, or the native stratum; Sino-Japanese,
which contains vocabulary from Chinese; Foreign, which contains more recent
technical and other words borrowed from foreign languages that are not Chinese;
and Mimetic, which contains the large number of words with sound-symbolism in
Japanese. As Itô and Mester (1995) explain, there are some phonological patterns
in Japanese, such as the voicing alternations of Rendaku, that hold in only a given

---

10 In fact, in some words with an initial /str/ cluster, [ʃ] can appear as a phonetic variant of /s/,
with pronunciations like [strɪt] and [ʃtrɪt] both being allowed for street (see, e.g., Durian [2007]).
Such a distribution, which appears to be allophonic on the basis of lexical identity, complicates
any attempts to say that [s] and [ʃ] are contrastive in this position on the basis of pairs like stick
and schtick.
morpheme class or classes – in the case of Rendaku, only in the Yamato class.\textsuperscript{11} Edwards and Beckman (2008: 23) similarly point out that “[t]he sequences . . . /ti/, /ʃe/, and /tʃe/ are attested only marginally, primarily in recent loan words from languages such as English.” Thus, each different stratum might be expected to have different relationships holding among particular sounds, or even different inventories of sounds.

In many cases the lexical exceptions to otherwise predictable relationships, such as those given in Section 3.1.2, are foreign or specialized words, as was the case with Bloomfield’s Menomini example described above. In other cases, as pointed out by Bye (2009), the marginal contrasts seem to be related to “expressiveness” or sound symbolism. For example, he points out that in North Saami, the aspirated rhotic /hr/ is primarily found in “verbs denoting sounds with a pronounced noise component, such as ‘sputter,’ ‘grate,’ or ‘neigh’” (Bye 2009: abstract). In still other cases, the specialized stratum may be due to some other sociolinguistic variable; Brown and Ogilvie (2008: 397), for example, claim that vowel length is “marginally phonemic” in French, but only in “careful educated speech.” Similarly, Chang (1998: 278) and Blust (2003: 19) note that mid vowels are marginally contrastive in Thao (an Austronesian language of Taiwan) and appear only in proper names and onomatopoetic words. Other examples include Yiddish ([l]/[l]); Auwera and König [1994: 396]), Cairene Arabic (the emphatics /ḷ/, /ṛ/, /ḅ/, and /ṃ/, as well as /p/, /v/ and /ʒ/; Watson [2002: 10]), Modern German (stress; Vennemann [1971: 110]), and Tunica (voicing; Moreton 2006).

3.3 Issues of variability and gradience

Another reason to declare a phonological relationship to be intermediate in some way is for there to be a high degree of variability in the way that it is produced, a criterion that seems to be an addition to the typical criteria described in Section 2.1. For example, Ladd (2006), in describing the relationship between [e] and [ɛ] in French and Italian, points out that some speakers do not make a distinction between the two or, if they do, have the distribution reversed from the standard variety. Similarly, Chitoran and Hualde (2007: 45) describe the distribution of diphthongs and hiatus in Spanish as being a “somewhat unstable” contrast as

\textsuperscript{11} Rendaku is the term for the voicing of the initial consonant of the second member of a compound or other multimorphemic sequence in Japanese; see, e.g., Vance (1987b: Ch. 10).
compared to that in other Romance languages, partially because many of the words that exceptionally have hiatus instead of a diphthong have it only optionally. Yliniemi’s (2005: 45) description of Denjongka of Sikkim, mentioned above, also appeals to variability as a contributing factor in the marginal nature of vowel length as a contrastive feature, saying that /y/ and /ø/ both tend to be predictably long or short based on the syllable structure that they appear in, but that “but both long and short /y/ and /ø/ appear in both open and closed syllables.” And Scobbie (2002: 7) mentions that [x] in Scottish English might be considered marginal in part because of its “propensity to merge with /k/ among many speakers.” Thus, there is a sense that variability in the realization of a sound is indicative of a lack of being “truly” contrastive in some sense, though how to measure such variability or the threshold at which a sound loses its contrastive nature is not at all defined.

Relatedly, phonetic gradience is also commonly associated with allophony, though it is more often presented as a consequence than as a diagnostic. For example, Cohn (1993) demonstrates that there is a phonetic difference between nasal vowels in English, where they are claimed to be allophonic variants of underlingly oral vowels, and nasal vowels in French, where they are claimed to be contrastive with oral vowels and underlingly specified as nasal. Specifically, Cohn shows that the nasal airflow in English nasal vowels increases (or decreases) gradually preceding (or following) a nasal consonant, while in French, there is a rapid transition between any oral and nasal segments and a high, sustained degree of nasal airflow throughout the nasal vowel. This is interpreted as being an artifact of the differing phonological specifications of the two. In English, where nasal vowels are allophonic, there is no need for a nasality target to be specified in the vowel, and nasal airflow is simply the consequence of interpolating between an unspecified vowel and a nasal consonant. In French, on the other hand, vowels are specified as being oral or nasal, and the timing and degree of nasal airflow is assumed to match those specified targets. Cohn uses this evidence to argue that nasalization in English is phonetic in nature while nasalization in French is phonological, but as Ernestus (2011) points out, gradience can be a characteristic of phonological representations, and it is often taken to be a sign of allophonic rather than contrastive properties (e.g., Giegerich [1992: 215] says that “many (perhaps all) allophonic processes” have a “gradient nature”). The extent to which gradience should be used as a criterion for determining allophony as opposed to contrast, and the exact nature of the gradience of any particular allophonic representation, are a matter of investigation.
3.4 Issues of frequency

A fourth general way in which phonological relationships can appear to be intermediate is when the elements within them occur with very low frequency. This scenario sometimes, of course, overlaps with the situations described above; for example, if a phone occurs only in foreign loanwords, it is likely to be less frequent, as well, and lexical exceptions tend to occur with lower frequencies than do regular patterns. As an example, Dresher and Zhang (2005: 33) say that “/y/ is a marginal phoneme in Oroqen [a northern Tungusic language of China], found only in a handful of words, many of which are borrowed from Chinese (Zhang 1996: 160).” As with variability, frequency seems to be an additional criterion that phonologists now think is relevant but which is not part of the traditional “set” of criteria described in Section 2.1.

For example, Watson (2002), in her description of Cairene Arabic, specifically emphasizes that many of the “marginal or quasi-phonemes” that she lists (see Section 3.2) are found in only a “few” words. Schreiber (1974: 749), also discussing Arabic, says that “strict adherence to the conditions of invariance and bi-uniqueness typically requires that (at least in the dialects) stress be considered (marginally) phonemic, given the existence of a usually quite small number of forms in which stress appears superficially to be unpredictable” (emphasis added). Adelaar and Himmelmann (2005: 398) describe the occurrence of final glottal stop in the Tandek dialect of Kimaragang as being “marginally contrastive” as compared to its “highly contrastive” status in the Pitas dialect, apparently coming to this conclusion on the basis of frequency of occurrence: they say that in Tandek, final open syllables (without glottals) are “quite rare,” though “a few minimal pairs can be found.” Similarly, Bals et al. (2007: 10), in discussing the inventory of diphthongs and triphthongs in North Saami, seem to appeal to frequency in describing “a marginal contrast between [au] and [aau] – generally, we find [aau], but we have also encountered njauge ‘smooth-coated dog’, raussa ‘baby diapers’.” Bye (2009: abstract) discusses a “marginal contrast” in certain varieties of RP between short and long /æ/, where “[l]ong /æ/ only occurs in a handful of highly frequent monosyllabic nominals ending in a voiced non-continuant, mostly adjectives like ‘sad’, while the short variant is used elsewhere.” Sohn (2008: 53) makes more explicit reference to frequency differences in describing the “marginally contrastive” status of [l] and [n] in word-initial position in Korean, specifically comparing counts of occurrences in various positions.

It should be noted that in both the issues of foreign borrowings (Section 3.2) and low frequency (Section 3.4), languages may end up with individual elements that are described as “marginal” (e.g., the presence of [x] in English or the
examples from Cairene Arabic given by Watson [2002]). This is of course a slightly different scenario from one in which a pair of sounds is thought to be “marginally contrastive,” but the two often go hand in hand. That is, one way in which the relationship between a pair of sounds can be marginal or otherwise intermediate is in the case in which one of the elements of that pair is rare or unusual, such that there is an asymmetry to the distribution of the elements. (For example, in English, the marginality of [x] might lead to a marginal contrast between [x] and [k], as in minimal pairs such as loch [x] vs. lock [k].) Goldsmith (1995) separates such asymmetry cases from other types of intermediate relationships due to issues with predictability of distribution or other factors; see further discussion in Section 4.4.

3.5 Subsets of natural classes

Another set of intermediate relationships can be defined by looking at the ways in which segments are divided into natural classes. For example, Austin (1988), in describing voicing contrasts in Australian aboriginal languages, distinguishes between “full” and “partial” contrasts at least partly based on how many members of a natural class show the contrast. As an example of a “full” contrast, he gives voicing in word-initial position in Murinypata: All stops contrast for voicing in word-initial position, so stop-voicing is a full contrast in this position. In other positions, there is only a “partial” stop-voicing contrast, because not all stops contrast – for example, after an alveolar stop, only bilabial stops contrast for voicing; after a velar stop, both bilabial stops and laminal stops contrast for voicing; after a tap, bilabial stops, laminal stops, and velar stops all contrast for voicing, but apical stops do not. Moreton (2006) also seems to make use of this kind of argument in explaining why voicing contrasts are “marginal” in both Woleaian and Chukchee. In each language, only one pair of segments illustrates a voicing contrast ([ʂ] vs. [ʐ] in Woleaian; [k] vs. [ɡ] in Chukchee); these are classified by Moreton as “marginal.” Similarly, Lowes (2006: 88) describes tone in Kurtoep as being “marginally contrastive” in that tones are contrastive on vowels after sonorant consonants but not entirely contrastive after obstruent consonants; this appears to be a case of ongoing tonogenesis.

A slightly different use of natural classes for determining “partial” contrast is found in Frisch et al.’s (2004) discussion of the Obligatory Contour Principle (OCP; “no adjacent identical elements”) in Arabic. In creating a similarity metric for measuring the strength of the OCP, they distinguish between features that are “fully contrastive” and those that are “partially contrastive”: Partially contrastive features are those that in some combinations form natural classes and in some
combinations do not. For example, they claim that [voice] is a partially contrastive feature because, for instance, the addition of [+voice] to the natural class [+continuant] creates a new, smaller natural class, but the addition of [+voice] to the natural class [+sonorant] does not change the membership of the class. They claim that partially contrastive features have less of an impact on similarity than do fully contrastive features.

3.6 Issues of phonetic similarity

Finally, a rather more tangential way in which differentiation is made between kinds of contrast is that some contrasts are more phonetically similar to one another than others. This is quite a different state of affairs than the previous issues, which are all clearly grounded in phonological issues of the role that elements play in the system, and which are true examples of intermediate relationships. Differences in phonetic similarity, on the other hand, are generally thought to be somewhat orthogonal to “how contrastive” pairs of sounds are and instead are presented as issues of “how similar” contrasts are (though cf. the use of phonetic similarity as a criterion). They are included here, however, because the terminology is sometimes similar, and it is important to recognize the different issues that are at stake.

One way of determining phonetic similarity is based on the number of features the elements of the contrast share. Jakobson (1990: 245) makes reference to “complete” versus “partial” contrasts, giving as an example of a “complete” contrast the difference between [ɪ] and [ŋ], which share no phonological features, and as an example of a “partial” contrast the difference between [p] and [t], which share all but one feature. He further divides partial contrasts by the number of differing features. For example, a difference of one feature is “minimal” while a difference of two features is “duple.” Campos-Astorkiza (2007) shows that minimal contrasts – those differing in exactly one property – are sometimes singled out by phonological processes, indicating that such distinctions are indeed meaningful. For example, she argues that in Lena, vowel harmony is triggered only by “inflectional vowels that are minimally contrastive for height” (Campos-Astorkiza 2007: 5). Such arguments depend, of course, on a phonological theory that makes use of features, and the extent of the contrast will be dependent on the way that features are assigned to segments. Interestingly, however, Connine et al. (1993) also show that pairs of sounds that are minimally contrastive in this sense tend to prime each other better than those that are maximally contrastive: e.g., *battern* acts as a lexical prime for *pattern*, while *rattern* does not.
A similar way in which differentiation is made between kinds of contrast is that some contrasts are considered to be “stronger” or “weaker” than others in terms of perceptual salience. For example, Steriade’s “P-map” approach (2001) encodes a map of the relative perceptibility of various pairs of sounds in various contexts; Steriade claims, for instance, that “voicing is, in any context, perceived as less distinctive than contrasts based on obstruency differences” (2001: 20). Thus, the contrast between [p] and [b] is said to be less salient than that between [b] and [m]. Such differences in salience have led to discussions of contrast “enhancement” or changes that make a particular contrast better (e.g., changing [fθinos] ‘cheap’ to [ftinos] in Greek to enhance the distinction between the first two consonants; see Hume and Johnson [2001]). In some sense, then, the contrast between [f] and [t] is thought to be better or stronger (i.e., more salient) than the one between [f] and [θ]; the grammar has moved in the direction of choosing the more salient distinction over the less salient one. As mentioned above, however, differences in salience are not generally thought to reflect any differences in phonological distinctiveness; for example, /f/ and /θ/ are assumed to be just as much distinct phonemes as /f/ and /t/ in Greek. Thus, even though the terminology of contrast enhancement is sometimes used as if it were the contrast itself that is becoming stronger, it is generally assumed that it is the cues to the contrast that are being enhanced, not the contrast itself. These distinctions based on perceptual factors, then, do not typically result in the kinds of intermediate phonological relationships that are otherwise under discussion in this paper, though they are commonly discussed in the literature on enhancement theory, dispersion theory, and salience. (And, as noted in Sections 2.2 and 3.1.2, phonetic similarity is often used as a filter on possible phonological relationships, and changes in phonetic similarity might thus be expected to change language users’ analysis of such relationships, as claimed in Janda [1999].)

3.7 Summary

In summary, there are a large number of instances in which the traditional binary distinction between “contrast” and predictable “allophony” is inadequate for describing the actual distribution of phonological entities in the world’s languages. This has been noted by other authors as well: Hualde (2005: 20) says that “there are areas of fuzziness probably in every language”; Ladd (2006: 14) claims that “instances of these problems are widely attested in the phonology of virtually every well-studied language”; and Scobbie and Stuart-Smith (2006: 15) state that, “in [their] experience . . . every language has a rump of potential / actual near-phonemes” (emphasis original).
In addition to the specific cases described above, there are many cases where terms indicative of intermediate relationships are used without further qualification. For example, Collins and Mees (1991: 85) mention in passing that short /a/ and long /a:/ are in a “quasi-allophonic” relationship in Welsh; Svantesson (2001: 159) claims that there is a “(marginal) contrast between dental [n] and alveolar [ŋ]” in his Southern Swedish dialect of Getinge; Fougeron et al. (2007: 1) state as fact that in French, “/a/ and /œ/ do not contrast and are in a quasi-complementary distribution”; Hildebrandt (2007: 4) says in a passing description of the Nepalese language Gurung that segment duration is “diachronically young and only marginally contrastive”; and Baković (2007: 17) mentions in a footnote that in Lithuanian, “[p]alatalization of consonants is automatic before front vowels and semi-contrastive otherwise.” None of these studies elaborates on the details of what makes these relationships intermediate, making it difficult to evaluate the possible effects of this intermediacy on the language in question.

There are also cases involving complex interactions of several of the types of intermediate relation listed above. For example, Crowley (1998) describes the complex relationship between [s] and [h] in Erromangan, an Oceanic language spoken on an island in Vanuatu in the southwest Pacific. Much of the time, [s] and [h] are in complementary distribution, but there are a few minimal pairs such as *esen* ‘ask for’ versus *ehen* ‘put in to,’ or *nmas* ‘large’ versus *nmah* ‘death.’ Additionally, words with s are often freely pronounced with h, though the reverse is not true. Such variation is common in medial and final position, but not in initial position. There are also diachronic, sociolinguistic, and religious factors that play into the distribution, as Crowley observes: “While it seemed initially that there was a possible phonemic contrast between s and h, in one of the few minimal pairs I had, the contrast was being maintained only about 40% of the time in the word *nmas* ‘big’. In the supposedly contrasting word *nmah* ‘death’, the contrast was maintained all the time, except that it was lost on Sunday mornings between 10.00 and 11.00 o’clock, or, on a bad day, 11.30. This, I should point out, is also only when singing, because when preaching and praying spontaneously in church, people were still coming out with the usual *nmah* for ‘death’, rather than *nmas*” (155). Crowley concludes that it is impossible to determine an either/or kind of relationship when one is faced with what he terms “mushy phonemes” (165).

Despite these complex cases, it is clear that there are a number of consistent problems in the application of criteria to the determination of phonological relationships. Issues of how to define predictability of distribution, what to do with substrata of a language, how to incorporate variability and frequency into an analysis, and how classes of sounds should be divided up all cause problems for a straightforward division of sounds into contrast or allophony. These issues are
further complicated by theory-internal issues about the representation of phonological categories and units. The next section briefly outlines the various approaches that have been used by phonologists for dealing with such a widespread amount of intermediacy.

4 How to handle intermediate phonological relationships

Although this is the first paper to lay out an extensive typology of intermediate phonological relationships, problems and shortcomings with the traditional distinction between contrast and allophony have been noted previously, and consideration has been given to the theoretical underpinnings of the definitions of contrast and allophony. Despite the fact that contrast is often still believed to be one of the central notions of phonological theory (e.g., Scobbie [2005: 8]: “[P]honology has the categorical phenomenon of contrast at its core”), a number of phonologists have questioned the traditional definitions. For example, as Steriade (2007: 140) points out in her article on contrast in the Cambridge Encyclopedia of Phonology, “[T]he very existence of a clear cut between contrastive and non-contrastive categories . . . in individual grammars” is contentious. This section provides an overview of the kinds of approaches to dealing with the problems posed by intermediate relationships that have been made in the literature. Most of the proposals have been focused on dealing with the problems associated with predictability of distribution (Section 3.1), but many of them also incorporate strategies for addressing the other areas of difficulty described above. Therefore, this section is organized by the type of approach rather than by the specific problems they are designed to address.

4.1 Polysystemicity: Different strata or subsystems

One frequent strategy for handling the existence of intermediate phonological relationships is to relegate the atypical patterns to separate parts of the grammar. This strategy is particularly common when there are patterns that are easily grouped together and stem from the same historical source, such as a group of words with exceptional phonological patterns that were all borrowed from the same source language; this approach is therefore most directly relevant for addressing the issues raised in Section 3.2. Fries and Pike (1949: 29–30) introduced the idea of “coexistent phoneme systems” to account for the numerous different
conflicts that arise between the native, “normal” phonology and various abnormal linguistic elements, such as borrowed or foreign words, interjections, “extra schoolroom contrasts,” or stylistically altered speech. They claim that trying to devise a unified system for all of these different types results in “internally inconsistent and self-contradictory analyses” (Fries and Pike 1949: 30). This result, however, seems to follow because they assume a binary choice (cf. the approaches in Section 4.4, in which gradience is allowed into the system): Either an exceptional form is ignored and only the rest of the phonological system is analyzed, or the exceptional form is accepted, wholesale, into the phonological system and any regularities that are therefore disturbed by its introduction are simply not considered regular anymore. It is obvious why neither of these solutions is satisfactory; the former ignores part of the linguistic system controlled by native speakers of a language, while the latter ignores regular, predictable patterns that hold over much of the language. Relegating exceptional forms to a more peripheral part of the grammar allows them neither to be ignored nor to interfere with the more regular patterns of the larger system.

This approach of having multiple systems has been adopted for many languages over the past sixty years,\textsuperscript{12} despite objections such as that of Bloch (1950: 87), who deems it “unacceptable” to try to separate out different parts of the “necessarily single . . . network of total relationships among all the sounds that occur in the dialect.” Itô and Mester (1995) review some languages that have different phonological strata and focus on describing the well-known case of Japanese, which, as described above, is traditionally assumed (with the notable exception of Bloch [1950]) to have four different strata. Recall that there are phonological patterns in Japanese, such as the voicing alternations of Rendaku, that hold in only a given morpheme class or classes – in the case of Rendaku, only in the Yamato class. However, it is not feasible to assume that each class has its own entirely separate phonology, because some patterns are found across multiple classes or even in all classes. Nor can one assume that the classes are nested hierarchically with all patterns holding for the innermost class, and fewer and fewer patterns holding toward the periphery, because there is no way to order the classes as being proper subsets of each other. Instead, Itô and Mester adopt a complex system of overlapping “constraint domains,” where each constraint on phonological representation is assumed to be applicable in certain parts of the lexicon, some of which are overlapping. Their account maintains the assumption of at least three separate lexical strata, though the non-homogenous character of

\textsuperscript{12} Note that Firthian Prosodic Analysis is another prominent theory that advocates for a poly-systemic approach, though it does not include the traditional notions of “phoneme” or “inventory” (see, e.g., Firth [1948], Ogden and Local [1994]).
the “Foreign” stratum forces a rejection of this class as a separate entity, at least phonologically.

One advantage to assuming this kind of stratified model is that the different strata do often reflect unified subgroups of the lexicon that are distinct from all other parts. As long as these are either closed classes or classes that can be entered only by items sharing the unifying characteristic (e.g., another word borrowed from the same foreign language), then such a separation of the phonology is certainly appropriate.

However, when phonological patterns from one stratum affect items from another stratum, or lexical items seem to cross over into different strata, it is less clear that having such dividing lines is the best analysis. For example, Itô and Mester (1995) describe a difference between “assimilated foreign words” that are subject to a phonological constraint against non-palatal coronals appearing before [i] (e.g., [cɕiːmu] ‘team’) and “unassimilated” foreign words where the constraint does not hold (e.g., [tiːn] ‘teen(ager)’). This distinction, which could be assumed to be a marker of “different strata,” is descriptive rather than following from any principled explanation; some foreign words are simply subject to the constraint, and some are not (as Itô and Mester point out). Also problematic is the observation that there are some native words that belong to what Itô and Mester call the periphery – the area of the grammar in which not all constraints hold. Thus, it is not the case that the peripheral area of the grammar corresponds with a particular stratum of the lexicon, and so stratification does not solve the problem of having conflicting phonological patterns. Instead, it seems as though in at least some non-fossilized areas of the grammar, certain phonological patterns simply hold to a greater or lesser degree over the entire lexicon. Bye (2009) notes similar problems with trying to use lexical strata to account for the marginal status of /hr/ in North Saami, where although most words containing /hr/ can be said to belong to some sort of sound symbolism stratum, there are three “regular” (i.e., non-symbolic) words containing /hr/. Bye concludes that in such a case, “it [is] impossible to define a lexical stratum” and that in general, “strata should be well-defined and have independent linguistic motivation” (13).

Furthermore, simply relegating some sections of the lexicon to a different phonological grammar does not really explain their behavior. While it might be true that a more peripheral section of the grammar is more prone to loss or assimilation, simply labelling it as peripheral does not explain why it has the properties it does (and it is clear that it is not just a case of loanwords belonging to the periphery, as mentioned above; see also Kreidler [2001: 448]). Thus, while there tend to be good historical reasons for adopting a stratified approach, care must be given to ensure that the stratification is done in such a way as to reflect some principled reality in the minds of speakers.
An approach somewhat similar to the idea of strata is mentioned in Anderson (2003: 269), who says that “[t]aking contrast seriously would [lead] to the postulation of polysystemicity, the recognition that in any language there is lexically not necessarily a single inventory of vowels and consonants, a single lexical inventory of phonemes.” That is, there could be a different inventory of possible phonemes in different positions in a language; in English, for example, there might be a different inventory of segments in the context [#s__V] than there is in the context [#__V]. Blust (2003) assumes this kind of multiplicity of inventories in his discussion of Thao phonology, saying that the language has two inventories of vowels, one of which is apparent only in proper names and onomatopoetic words. This multiple-inventory approach is similar to the notion of strata in that there is not a single grammatical system involved, but it stratifies the grammar at the level of the phoneme and its conditioning environments rather than the level of the word. While it is a potentially useful concept for understanding intermediate relationships that are the result of contextual neutralization, then, it does not account for account for intermediacy based on frequency or specialization of word classes, for example.

Another approach that fits in to this category of polysystemicity is the concept of “positional” markedness or faithfulness, in which different positions in the grammar are subject to different grammatical constraints (e.g., Selkirk 1994; Beckman 1997, 1998; McCarthy and Prince 1995; Casali 1996; Smith 2001, 2008; Steriade 1997; Zoll 1997). Beckman (1998: 1), for example, argues that certain positions such as “roots, root-initial syllables, stressed syllables, and syllable onsets” are perceptually prominent or “strong” and so may allow contrasts to occur in them where other positions disallow them. Smith (2001) further claims that it is not just phonologically defined positions that can be strong, illustrating that nouns act as “strong” positions – nouns in Spanish, for example, have phonemically distinctive stress, but the stress patterns of verbs in Spanish is predictable. On the other hand, it has also been argued (e.g., Steriade 1997; Zoll 1997) that some positions (e.g., unstressed syllables, syllable- or root-final positions, etc.) are particularly “weak” and therefore license fewer contrasts than other, stronger positions. Both such approaches focus on the positions rather than the contrasts themselves, though the effects may be similar: a particular contrast may be enhanced in a strong position or neutralized in a weak position, in comparison to other contrasts, creating disparity among the contrasts in a language. Interestingly, these approaches might be fairly straightforwardly extended to account not just for positional effects but also effects due to frequency, foreignness, and the like, assuming that constraints on contrasts can be indexed to such factors (e.g., one could imagine positing a set of constraints mandating faithfulness in foreign words as opposed to those mandating faithfulness in native words, and ranking
them as $\text{FaithForeign} \gg *[x] \gg \text{FaithNative}$ in order to get different distributional patterns in the two different strata. At a certain point, however, such treatments tend to turn into mere descriptions of the facts of the language unless one is careful to provide external motivation for such constraints and evidence that the classification of certain words as “foreign” or otherwise special is psychologically real.

### 4.2 Functional approaches

Another approach to dealing with different kinds of contrastive relationship is to quantify a relationship’s function in the greater system of its language. The most typical way of doing so is to calculate a relationship’s *functional load* in a language (e.g., Martinet 1955; Hockett 1955, 1966; Surendran and Niyogi 2003; Wedel and Branchaw 2011). The functional load of a contrast can be roughly estimated by counting the number of minimal pairs that instantiate that contrast, with a larger number of minimal pairs indicating a higher functional load. More specifically, functional load is usually defined in terms of information loss (cf. Shannon and Weaver 1949): If there is a contrast between two sounds, A and B, in a language, how much would the entropy (uncertainty) of the language change if the contrast between A and B were to disappear? Functional load has been predicted to correlate with resistance to sound change, in that contrasts with a higher functional load have been predicted to resist sound change (e.g., neutralization) more than contrasts with a lower functional load (e.g., Martinet [1955], though cf. King [1967]). As Wedel and Branchaw (2011) point out, however, functional load is only one possible influence on sound change; rather than looking at any one case study, Wedel and Branchaw build a database of “recent mergers in languages for which there exist frequency-coded phonemic word lists” and show that, statistically speaking, across languages, “occurring mergers reduce word-level entropy less than non-occurring mergers,” in line with the functional load hypothesis. Such approaches indicate that the amount of work a contrast does in a language does seem to have consequences for language use, such that not all contrasts should be considered the same. At the same time, functional load is a measure of how much work a particular contrast does in a language as compared to other contrasts and does not account for other aspects of intermediacy – for example, functional load does not distinguish between pairs of sounds with different degrees of predictability of distribution. That is, a particular pair of sounds might have a high functional load in two languages, but be only partially unpredictably distributed in one language and highly unpredictably distributed in the other. This is not to say that functional load is not a useful measure, but rather to point
out that using it does not account for all of the kinds of intermediate relationship described above. In some cases, authors have even appealed to functional load as evidence for the marginality of a contrast; Gleason (1961: 297–298), for example, claims that “downstep has a somewhat different status in the phonemic system of Nguni from most other contrasts. Its functional load is low; minimal pairs are few and seldom commute in usable contexts,” and suggests that it be considered a “hemiphoneme.” Lowes (2006) similarly describes tonogenesis in Kurtoep as showing contrasts with increasing functional loads that are currently in a state of marginal contrast.

Hume and Mailhot (2011) extend the notion of functional load to individual features, illustrating that the set of features in a given phonological system can be differentiated by how much work each feature does in distinguishing segments in the inventory from one another. For example, in a toy language with five vowels and the features [vocalic], [high], [labial], [coronal], and [delayed release], they show that the feature [high], which minimally distinguishes two pairs of vowels in the system ([i]/[e] and [u]/[o]), does more work than the features [labial] and [coronal], which each minimally distinguish one pair of vowels in the system ([o]/[a] and [e]/[a], respectively), which in turn do more work than [vocalic] and [delayed release] which distinguish none of the vowels in the system. Hume and Mailhot quantify this difference in terms of entropic contribution (cf. Shannon and Weaver 1949). Thus, they provide a means of quantifying the degree to which a particular feature is contrastive or distinctive within a particular system.

Another functional approach more generally is found in Ussishkin and Wedel (2002, 2009), who describe what they call “effective contrast,” a measure of how distinct a particular lexical item is from its lexical neighbors in terms of the efficiency with which it can be accessed, as a function of (i) neighborhood density, (ii) the frequency of occurrence of the particular lexical item, and (iii) the degree to which the lexical item participates in allomorphic alternations. This, then, is not a measure of the degree to which two sounds are distinctive but rather a measure of how different words are from each other from the point of view of a language user trying to access those words; the relationships between individual sounds are dependent, then, on the words in which they are embedded. Ussishkin and Wedel show the results of a number of case studies that suggest that effective contrast is correlated with the structure of lexicons in the world’s languages. They do not, however, provide any explicit way of calculating effective contrast as a function of the three factors they think are relevant.

All of the functional approaches described above can be used to address issues of frequency such as those discussed in Section 3.4. In particular, sounds that are infrequent are likely to be part of relationships with a low functional load
and have a lower effective contrast (though the latter will be mediated by the
density of the neighborhood of words the sounds occur in). Such approaches
are also based on the functionality of contrasts in the overall linguistic system,
which could be argued to touch on issues of the psychological reality of contrasts.
At the same time, such approaches so far either do not take into account other
criteria for determining contrast (functional load) or do not provide an explicit
means of balancing the various criteria that they do take into account (effective
contrast).

4.3 Enhanced machinery and representations

An alternative method of dealing with intermediate relationships is to enhance
theoretical devices and representations in some way. This section discusses
four approaches to changing phonological machinery. The first, presented in
Kager (2008), is an Optimality-Theoretic approach that very specifically ad-
dresses problems with the notion of predictability of distribution (Section
3.1). The second is Inkelas’ (1995) archiphonemic underspecification, designed
to account for the same kind of phenomenon as Kager (2008), though not
in the same context. The third, presented in Ladd (2006), relies on less theory-
specific tools, appealing to general patterns of cognitive organization, and is
intended to cover a number of the issues described in Section 3. The fourth is
the notion of paradigm uniformity, widely discussed by many phonologists,
though not generally for the purpose of accounting specifically for intermediate
relationships.

As mentioned in Section 1.2, Optimality Theory already distinguishes among
four types of phonological relationship (full contrast, complete neutralization,
partial neutralization, and allophony). Kager (2008) argues that at least a fifth
type of relationship should be included and describes an Optimality-Theoretic
approach to lexical irregularities in which one set of words in the lexicon under-
goes alternation, while other sets, which contain each of the alternants, do not.
He terms this kind of situation *neutrast* – a combination of “neutralization” (in
the alternating sets) and “contrast” (in the non-alternating sets) – and explains
that, like full contrast, contextual neutralization, and allophony, this is a type of
distribution of segments that must be accounted for. As an example, consider the
distribution of short and long vowels in Dutch, shown in (9); some stems always
contain a short vowel as in (9a), some always contain a long vowel as in (9b), and
some alternate between the two as in (9c).
(9) Distribution of short and long vowels in Dutch (Kager 2008: 21):

a. Non-alternating short vowel (many stems):
   - kl[ɑ]s ~ kl[ɑ]sen 'class(es)'
   - p[o]t ~ p[o]ten 'pot(s)'
   - h[e]g ~ h[e]gen 'hedge(s)'
   - k[i]p ~ k[i]pen 'chicken(s)'

b. Non-alternating long vowel (many stems):
   - b[aː]s ~ b[aː]zen 'boss(es)'
   - p[oː]t ~ p[oː]ten 'paw(s)'
   - r[eː]p ~ r[eː]pen 'bar(s)'

c. Alternating short-long vowel (few stems):
   - gl[ɑ]s ~ gl[aː]zen 'glass(es)'
   - sl[o]t ~ sl[oː]ten 'lock(s)'
   - w[e]g ~ w[eː]gen 'road(s)'
   - sch[i]p ~ sch[eː]pen 'ship(s)'

Kager (2008) proposes a system of “lexical allomorphy,” in which a single lexical item can have more than one lexical entry; the lexical entry for the stem ‘glass’ therefore would have both gl/ɑ/z- and gl/aː/z-. Although the grammar will force any input representation into a grammatically acceptable and optimal output, as is always the case in OT, the presence of multiple inputs means that there can be multiple output forms, as well. For non-alternating stems, highly ranked faithfulness constraints force the non-alternation; for alternating stems, faithfulness is always satisfied\(^\text{13}\) (because there are two possible inputs), and so markedness constraints determine the optimal alternant. Kager also relies on Output-Output faithfulness constraints to rule out having extraneous pairs of alternating stems – any alternating stem must be the result of re-ranking an OO-Faith constraint fairly low in the hierarchy.

Under Kager’s (2008) account of the typology of contrast, there are four basic types of constraints (two faithfulness, one Input-Output (IO-Faith) and one Output-Output (OO-Faith), and two markedness, one specific (MS) and one general (MG)), which result in six basic types of distributions, shown below in (10). (In (10), the subscript \(G\) refers to the form a word takes in the general case, while the subscript \(S\) refers to the form in the specific case. [αF] and [-αF] refer to the feature specification of the given class of words in the given environment and represent the features that are potentially contrastive. Each of the three columns represents a word class that contains words that have, at least potentially, the

---

\(^{13}\) Or perhaps faithfulness is never satisfied, depending on one’s perspective.
same phonological contexts as words in the other classes. Words in the left column have \([\alpha F]\) in both the general and specific context; words in the right column have \([-\alpha F]\) in both the general and specific context; and words in the center column have \([\alpha F]\) in the general context but \([-\alpha F]\) in the specific context. For example, a system with full contrast between \([\alpha F]\) and \([-\alpha F]\) allows both \([\alpha F]\) and \([-\alpha F]\) to occur in all contexts and so has words in both the left and right columns.)

(10) Factorial Typology of Allomorphy (Kager 2008: 33):
   a. Neutrast: IO-Faith » MS » MG, OO-Faith
      \([\alpha F]_g \sim [\alpha F]_s \sim [-\alpha F]_s \sim [-\alpha F]_s\)
   b. Full contrast: IO-Faith, OO-Faith » MG, MS
      \([\alpha F]_g \sim [\alpha F]_s \sim [-\alpha F]_g \sim [-\alpha F]_s\)
   c. Contextual neutralization: MS » IO-Faith » MG, OO-Faith
      \([\alpha F]_g \sim [-\alpha F]_s \sim [-\alpha F]_g \sim [-\alpha F]_s\)
   d. Total neutralization I: MG, OO-Faith » IO-Faith, MS
      \([\alpha F]_g \sim [\alpha F]_s\)
   e. Total neutralization II: MS, OO-Faith » IO-Faith, MG
      \([-\alpha F]_g \sim [-\alpha F]_s\)
   f. Complementary distribution: MS » MG » IO-Faith, OO-Faith
      \([\alpha F]_g \sim [-\alpha F]_s\)

By adding both lexical allomorphy and OO-Faithfulness constraints, Kager’s (2008) approach allows for more levels of distribution than the standard OT approach, which predicts only types b, c, d, and f of (10). These additions increase the generative power of an OT account, and in doing so, provide a formal account of neutrast situations. One potential problem, however, is that this approach does not allow for differences within a given level. Specifically, type c, contextual neutralization (which Kager also refers to as “partial contrast”) still encompasses most of the different scenarios described in Section 3.1. There is no way to capture the difference between cases that are mostly predictable, but with a certain degree of contrast, and cases that are mostly contrastive, with a certain degree of predictability.

To take a concrete example of a case in which such differences seem at least intuitively important to distinguish, consider the case of a Japanese contrast that is mostly predictable (as compared to the Dutch case, where the contrast is mostly unpredictable). As described in Section 4.1, in the Yamato, Sino-Japanese, and Mimetic strata of Japanese, the sequence [ti] does not occur; when it would arise through, for example, suffixation, a palatal coronal appears instead: [cɕi] (e.g., [kat-e] ‘win (imperative)’ vs. [kacɕ-i] ‘to win’). In some foreign words, this generalization holds and palatalization occurs (e.g., [cɕi:mu]
'team'), while in others, it does not apply, and the non-palatal surfaces (e.g., [ti:n] ‘teen(ager)’). Thus, with the advent of foreign borrowings, the difference is no longer phonologically conditioned, and Kager’s style of analysis would be appropriate. According to this approach, the way to encode partial predictability is through high-ranking specific markedness constraints. Kager also specifies that all constraints are universal and there are no morpheme-specific constraint rankings. To analyze the Japanese case, then, which is an example of neutrast, there must be a (universal) markedness constraint, *[ti], that penalizes [ti] sequences, along with a faithfulness constraint, e.g., Faith(Pal), that penalizes changes in palatalization between the input and the output. To achieve the variation in loanwords, it must simply be the case that /t/ and /ɕ/ are contrastive in Japanese, and the difference in the outputs is guaranteed by faithfulness to differing input forms, as shown in Table 1. The alternating forms are generated through lexical allomorphy; each has two input forms, allowing the lower-ranked markedness constraints to select the appropriate input.

This solution has a number of potential drawbacks. First, all native stems that alternate are subject to lexical allomorphy (e.g., the lexical entry for ‘win’ is /kat/~/kacɕ/). Introducing lexical allomorphy for all the native alternating words,

### Table 1: Tableaux for the neutrast of [t] and [ɕ] in Japanese

<table>
<thead>
<tr>
<th></th>
<th>/kat/ or /kacɕ/ + /e/</th>
<th>Faith(Pal)</th>
<th>*[ti]</th>
<th>*[ɕe]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>kate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>kacɕe</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>b</td>
<td>/kat/ or /kacɕ/ + /i/</td>
<td>Faith(Pal)</td>
<td>*[ti]</td>
<td>*[ɕe]</td>
</tr>
<tr>
<td></td>
<td>kati</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td></td>
<td>kacɕi</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>/ɕiːmu/</td>
<td>Faith(Pal)</td>
<td>*[ti]</td>
<td>*[ɕe]</td>
</tr>
<tr>
<td></td>
<td>tiːmu</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td></td>
<td>ɕiːmu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>/tiːn/</td>
<td>Faith(Pal)</td>
<td>*[ti]</td>
<td>*[ɕe]</td>
</tr>
<tr>
<td></td>
<td>tiːn</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td></td>
<td>ɕiːn</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
however, means that the introduction of a few non-native contrasts entirely restructures a large part of the native lexicon. Furthermore, this restructuring introduces a rather arbitrary redundancy in that all the forms that alternate happen to have input forms with [t] and [cɕ]. The generalization that the sequence [ti] is dispreferred in favor of palatalization before [i] is relegated to a coincidence within a large set of lexical items.

A second problem is that this analysis gives preference to the small minority of forms that actually show the contrast in Japanese, rather than the vast majority that show the allophony. The examples of neutrast in Kager (2008) are ones in which the contrastive word classes predominate, and there are only a few alternating examples, making the appeal to lexical allomorphy less costly.

Third, the alternations in the native word ‘win’ in Table 1 are governed by markedness constraints; in addition to the constraint against [ti] sequences, there is also a constraint against [cɕe] sequences in Japanese. For this example, the combination of these two constraints correctly selects the output forms [kate] and [kacɕi]. But consider the form [katanai], the negative form of the verb. There is no particular evidence for a markedness constraint against [ta] or [cca] (both are in fact real native words of Japanese). Given the lexical allomorphy between /kat/ and /kacc/, both the output forms [katanai] and [kacɕanai] should be possible; only the former, however, is actually found.\footnote{It is possible that a constraint against [cca] is simply a part of the universal markedness constraints that must, by assumption, be a part of the grammar of Japanese; it becomes apparent only once the native alternating words are subject to lexical allomorphy after the introduction of foreign words.}

As a general proposition, traditional phonological accounts of the marginal contrasts described in Section 3.1 rely on an analysis, like that of Kager (2008), in which the distribution is assumed to be basically contrastive, with the partial predictability of distribution being accidental. As seen by the above example, for cases in which the vast majority of forms alternate, this style of solution is unsatisfying.

Interestingly, the phenomenon of neutrast described by Kager (2008) is also accounted for by archiphonemic underspecification (e.g., Inkelas 1995, Inkelas et al. 1997). In this approach, rather than positing lexical allomorphy, it is assumed that stems that alternate are underspecified for the alternating feature. The most typical example is that of Turkish voicing alternations, in which some stem-final consonants are always voiceless, some always voiced, and some alternate between being voiced between sonorants and voiceless finally; thus, the consonants with voicing alternations would be underspecified for voicing. High-ranking input-output faithfulness constraints account for the non-alternating
stems; all alternating stems would violate such constraints equally, so that lower-ranked markedness constraints could govern the alternations. The archiphonemic underspecification account falls upon similar problems as the lexical allomorphy / OO-faithfulness approach. Specifically, in a case like that of Japanese described above, the introduction of a few new borrowings fundamentally restructures the entire lexicon. Before the borrowings, stems would presumably have been subject to markedness constraints and could be specified for palatalization underlyingly (and would be so specified under lexicon optimization). These native stems would then have to somehow lose their specification once (the much more limited number of) non-alternating stems were introduced to the language.

Another way of enhancing the phonology in order to account for intermediate phonological relationships is given by Ladd (2006). Ladd proposes a system of categories and subcategories, saying that “phenomena of stable partial similarity or quasi-contrast can be accommodated in a theory of surface representations if we assume that, like any other system of cognitive categories, phonetic taxonomy can involve multiple levels of organization and/or meaningful within-category distinctions of various kinds” (18). Thus, for example, one might have a super-category of vowels, within which are categories for A and E; within the E category, one might have both e and ɛ; and within the e category, one might have both [e] and [e:]. As Ladd (2006) shows is the case for French and Italian, these two phones seem to be less predictable than “true” allophones (because there are minimal pairs) but more predictable than “true” phonemes (because the contrast is neutralized in some environments, there is variability...

![Vowels](image.png)

**Fig. 1:** Example of Ladd’s (2006) category/subcategory approach to quasi-contrast
among speakers, and in some words the two are in free variation). Ladd does not specify how many levels are possible, but there is nothing in the argumentation to suggest that there could not be as many levels as there are phones in the language. One advantage to this approach, as Ladd points out, is that there is nothing inherently unstable about this hierarchical arrangement of categories. That is, it is quite possible to have a persistent quasi-contrastive relationship, without assuming that it is merely an intermediate stage between more stable situations of pure contrast or pure allophony. (Note too the similarity between this approach and one of archiphonemic underspecification; while Ladd focuses on the fact that there might be a level between the phoneme and the allophone, one could easily imagine an analogous level above the phoneme for the archiphoneme.)

While the approach in Ladd (2006) is intuitive and captures the “apparent closeness” between [e] and [ɛ] in French (Trubetzkoy 1969 [1969]: 78), it is unfortunately not yet fleshed out enough to be implemented as a practical matter. For example, Ladd does not specify how to decide which phones go in which level, or whether all nodes at the same level should be expected to behave the same way.

Ladd (2006) claims that Trubetzkoy’s argument, that the “closeness” stems from the neutralization of contrast, is inadequate (and thus presumably not the means by which pairs are assigned to levels), because not all neutralized contrasts show the same pattern. An example given is [t] and [d] in American English, which are neutralized to [ɾ] in trochees; Ladd says that unlike French [e] and [ɛ], there is no special relationship between [t] and [d] in English.15 The implication is that [t] and [d] should be at the top, “fully contrastive” level of the consonant hierarchy. This placement is problematic because it would mean that there is no indication in the model that the two are neutralized in some environments (without the addition of other rules, etc.). Furthermore, it is not clear how the analyst is supposed to know whether there is a “special closeness” between phones. Hume and Johnson (2003), for example, classify all neutralized contrasts as “partial contrasts” and show that, at least for the case of Mandarin tones, neutralization does affect the perceived similarity between tones. Ladd’s system does not provide guidelines for how to distinguish among different kinds of neutralizations of partial contrasts.

Furthermore, Ladd (2006) argues against the neutralization hypothesis because not all examples of marginal contrasts are related to neutralization – for instance, the examples in Section 3.1.3 are ones where phones are perfectly pre-

15 Note that the types of neutralization here are somewhat different: English [t] and [d] are neutralized to a third segment, [ɾ], while French [e] and [ɛ] are neutralized to something that is phonetically “indeterminate” between [e] and [ɛ], according to Ladd (2006).
dictable, as long as one is given access to non-phonological information. The implication by Ladd is that these cases should be included in the categorization/subcategorization system, which on the one hand is advantageous in that it presents a unified approach to intermediate relationships, but on the other is problematic in that it conflates very different sources of marginality that have not yet been shown to pattern the same way. Do we in fact want to put, for example, Scottish [aɪ] and [ʌɪ] into the same category as French [e] and [ɛ]? This remains an open question.

Another potential problem with the solution in Ladd (2006) is discussed by Hualde (2005: 20). Hualde points out that the boundaries that define categories such as the ones that make up the hierarchy in Ladd must have “fuzzy” boundaries: Although “phonological categories ‘tend’ to be discrete,” “the ranges of [particular phonetic elements may] show greater or lesser overlap depending on the dialect, the style and the speaker. The extent of the overlap may determine their categorization for a given speaker.” Thus, while the basic premise of different layers of phonological closeness may be exactly on track, the details of its implementation need to be developed, and, in particular, need to leave room for a wide range of disparate phenomena.

A fourth way of enhancing the phonological machinery that is quite well-known but not generally thought of as a way to account for intermediate relationships (though cf. Bye [2009]) is the concept of paradigm uniformity. Paradigm uniformity allows certain anomalous patterns to occur in some words on the basis of matching other forms of the word. For example, the apparently anomalous pattern of [ʌɪ] occurring in tide while [ɑɪ] occurs in tied in Scottish English (see Section 3.1.3) could be explained on the basis that there is a regular pattern of [ɑɪ] occurring word-finally, and the vowel in tied should match its uninflected counterpart tie. There are a number of ways in which paradigm uniformity can be implemented, from cyclic derivations to various kinds of correspondence constraints (e.g., Kiparsky 1982, Cole 1995, Kager 1999, Steriade 2000; Kenstowicz 2002). Such an approach, of course, is useful only for those cases of intermediate relationships in which some sort of analogical comparison is possible; it does not, for example, account for cases of intermediacy in which one sound is simply low frequency, or occurs only in a certain kind of word class (borrowings, sound symbolism, high frequency, etc.).

4.4 Gradience

Another kind of proposal for integrating intermediate relationships into a language’s phonology incorporates gradience into the description of phonological
categories. Building on the increasingly well-accepted assumption that linguistic phenomena are built on a “statistical foundation” (Scobbie 2005: 25), a number of phonologists have suggested that phonological relationships should also be considered in a statistical manner. To a certain extent, this is not incompatible with some of the other strategies for accounting for intermediate relationships given in Sections 4.1 and 4.3; having a number of different nesting strata or category levels moves the representations toward a more gradient effect, while maintaining discrete categories. It has been suggested, however, that a non-discrete, continuous model of phonological relationships is needed (see, e.g., Boulenger et al. [2011] and Ferragne et al. [2011], who give psycholinguistic arguments for their “Gradient Phonemicity Hypothesis”).

Goldsmith (1995), for example, suggests that there is a “cline” of contrast. In this model, phonological relationships are a reflection of the opposing pressures from the grammar on the one hand and the lexicon on the other. At one end of the cline, the distribution of two sounds is entirely governed by the lexicon (i.e., there is perfect contrast), while at the other end, the distribution is entirely governed by the grammar (i.e., there is perfect allophony). The model is thus predicated on the assumption that the grammar supplies all predictable information, while the lexicon is a repository for all unpredictable information (cf. theories of underspecification mentioned in Section 1.2). In between these two extremes, there are “at least three sorts of cases” (10), with the implication that there could be an infinite number depending on how the opposing forces are quantified. The points Goldsmith suggests are on this cline are given in (11).

(11) “Cline of Contrast” (Goldsmith 1995: 10–11):
   a. *Contrastive segments*: Two segments, x and y, can be found in exactly the same environments, but signal a lexical difference.
   b. *Modest asymmetry*: Two segments, x and y, are basically contrastive, but there is “at least one context” in which x is, for example, vastly more common than y.
   c. *Not-yet-integrated semi-contrasts*: Two segments, x and y, are contrastive in many environments, but there is “a particular environment” in which, for example, x is very common but y occurs only “in small numbers in words that are recent and transparent borrowings.”
   d. *Just barely contrastive*: Two segments, x and y, are basically in complementary distribution, but there is at least one context in which they contrast.
   e. *Allophones in complementary distribution*: Two segments, x and y, appear always in complementary environments.
Goldsmith’s (1995) proposal clearly incorporates many of the aspects of intermediate relationships described in Section 3, such as predictability of distribution, native versus foreign origins, and frequency of occurrence. There is, however, a certain degree of indeterminacy in assigning pairs of segments to a position on the cline. For example, the “modest asymmetry” case and the “not-yet-integrated semi-contrast” case are theoretically very similar, given that both are cases in which the lexicon plays a greater role in determining the relationship than does the grammar. The difference between the two cases is the source of the exceptions rather than the type of exceptions. One might ask why the fact that x and y marginally contrast in borrowings should mean that they are placed closer to the “grammatically conditioned” end of the scale. Furthermore, it is unclear how to detect the difference between cases that are basically contrastive with some predictability and those that are basically predictable with some contrastiveness. The assumption of a gradient cline, however, indicates that it could in theory be possible for segments to fall at any point along the scale, assuming there were a way to quantify the tension between grammar and lexicon. Bermúdez-Otero (2007), in a summary article “Diachronic Phonology,” accepts Goldsmith’s view of marginal contrasts as a useful addition to “classical” accounts of lexical diffusion based on evidence like that of Labov’s (1994) /æ/-tensing data. Bermúdez-Otero hints at the need for making the cline more quantitatively concrete, however, giving particular percentages of word classes that show or fail to show the expected tensing patterns.

Exemplar models have also been invoked as offering a possible solution to the problem of intermediate relationships. Such models assume that the details of all (or at least most) encountered speech are stored, and linguistic generalizations are emergent from the individual exemplars. Because individual tokens are stored in memory, and generalizations are emergent, these generalizations can reflect frequency information and give a fine-grained picture of the degree of overlap among categories. Scobbie and Stuart-Smith (2008: 108) explain that “[t]he exemplar view, though as yet very sketchy and lacking in many firm predictions, offers a clear mechanism for expressing gradual phonologisation, gradient contrast, nondeterminism, and fuzzy boundaries, all of which are real and pervasive in any phonology.” Specifically, exemplar models might account for intermediate relationships by encoding all sorts of specific information and nuanced generalizations; for example, that a particular sound is associated with foreign words, high- or low-frequency words, particular morphological categories, etc., or that particular sounds have a very specific kind of distributional pattern.

While details of the exemplar-based approach to phonological relationships remain to be worked out, many researchers view it as a promising approach. For example, Hualde (2005: 21) concludes that “[l]anguage is probabilistic (Bod
et al. 2003) and linguistic categories are emerging entities (Bybee 2001[b])," strongly suggesting an exemplar-based approach. Bermúdez-Otero (2007: 515) also claims to find at least a hybrid phonetic-exemplar-plus-phonological-encoding approach, along the lines of Pierrehumbert (2002, 2006), to be "worth pursuing."

Another model of intermediate relationships that is based on the notion of gradience is Hall (2009, 2012)'s Probabilistic Phonological Relationship Model (PPRM). This model is specifically designed to address intermediate relationships that are based on intermediate levels of predictability of distribution. Relationships are assumed to fall along a continuum of predictability, rather than being split into simple allophonic (predictable) and contrastive (unpredictable) categories. The model involves calculating the precise degree of uncertainty of choice between two sounds in a language, using the information-theoretic concept of entropy; sounds for which there is a high degree of uncertainty are more contrastive, while those for which there is a low degree of uncertainty are more allophonic. This approach is promising in that it provides a descriptive tool for cataloguing the various degrees of unpredictability that arise in the world's languages and describing the changes that relationships may undergo over time. In terms of practical use, the model depends on calculating distributions from allophonically tagged corpora of languages, which are somewhat difficult to come by, but it may be possible to extend the PPRM to experimentally derived data (see, e.g., Thakur [2010]). A somewhat similar approach is presented in Peperkamp et al. (2006), who propose a statistical learning algorithm, also based on information-theoretic concepts, to model the acquisition of allophonic relationships from corpora; their learning algorithm is not directly designed to model intermediate phonological relationships, but it does so inherently by treating predictability of distribution as a gradient notion. Note that both these approaches focus solely on cases of intermediate degrees of predictability of distribution, though their reliance on information-theoretic measures of probability means that at least issues of frequency can also be incorporated into the models.

In sum, there are a number of different strategies for incorporating gradience into phonological representations and thereby having a mechanism for describing phonological relationships that are not in fact binarily distinct themselves. Each of these strategies has some interesting insights into the issue and present promising avenues for future research, though none have been fully worked out for the full range of cases at this point.
5 Summary and conclusions

In summary, while the notions of contrast and allophony are commonly taken for granted as an important component of phonological analysis in almost any theory in the phonological literature, and while they are often assumed to be fairly straightforwardly determined by the criteria of predictability of distribution, lexical distinction, native speaker intuition, alternations, phonetic similarity, orthography, and/or parallelism to other patterns in the system, it should be clear that these criteria are in fact insufficient in their current forms to account for the wide range of phonological relationships actually seen in the world’s languages. The criteria may conflict with each other; they may be vaguely defined and thus difficult to apply; and they may not encompass all of the factors that phonologists deem relevant to the determination of relationships between sounds. When describing the relationships that hold among elements in a language, researchers should be explicit in their descriptions of any intermediate relationships, so that the nature and consequences of such relationships can be more fully understood. While there are a number of promising approaches for dealing with many of these problems, there is as yet no solution to all of them or even an entirely satisfactory solution to any one of them. It is my hope that the present typology of the problems we face will lead to the building of phonological grammars that can encompass all phonological relationships.

References


Bals, Berit Anne, David Odden & Curt Rice. 2007. The phonology of gradation in North Saami. Columbus, OH.


Chao, Yuen-Ren. 1957. [1934]. The non-uniqueness of phonemic solutions of phonetic systems. In Martin Joos (ed.), *Readings in Linguistics I: The Development of


Kochetov, Alexei. 2008a. Phonology and phonetics of loanword adaptation: Russian place 
Kochetov, Alexei. 2008b. Self-organization through misperception: secondary articulation and 
vowel contrasts in language inventories. In Peter Avery, B. Elan Dresher & Keren Rice 
Gruyter.
Routledge.
MA: Blackwell.
Phonetics, phonology, and sound change. Berlin: Mouton de Gruyter.
Gruyter.
(eds.), Experimental approaches to linguistics, 13–44. Oxford/New York: Oxford University 
Press.
Typology of Tone and Intonation Europe 2, Berlin, September 9.
Lu, Yu-an. 2012. The role of alternation in phonological relationships: Stony Brook University 
Doctoral dissertation.
Mackenzie, Sara. 2005. Similarity and contrast in consonant harmony systems. Toronto 
Mahanta, Shakuntala. 2008. Directionality and locality in vowel harmony: With special 
Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), The Blackwell companion 
Mon-Khmer Studies 27. 171–213.
Sino-Tibetan reconstruction: University of California Press.
Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), University of Massachusetts


Tadmor, Uri. 2002. Language contact and the homeland of Malay. Jakarta Field Station of the MPI-EVA.


