

## Chapter 3: Theories of syllable structure

### 3.1. Defining the syllable

The phonetic definition of the syllable is notoriously difficult. A common view is that a syllable is a prominence peak, but the definition says little about where syllable boundaries are. Also, it is unclear why some phonetic peaks are not treated as syllables, such as the [s] in *stop*, *extra*, and *cats*. Another common definition is that a syllable is related to a chest pulse, or a pulse of air pressure. However, this definition again says little about where syllable boundaries are. In addition, as Gimson (1970) puts it, it is doubtful whether a double chest pulse will be evident in a word like *seeing* [si:ɪŋ]. Nor can the pulse theory decide whether a word like *beer* [biə] in American English is one or two syllables.

Such problems have led to certain skepticism on whether syllables are real linguistic units. For example, Chomsky and Halle (1968) do not consider the syllable to be a relevant phonological entity. Similar reservations are expressed by Gimson (1970), Steriade (1999), and Blevins (2003).

Nevertheless, in some cases it is fairly clear what syllables are. For example, all people agree that *Canada* has three syllables and *America* has four, and many people agree that *after* can be divided as *af-ter*, *cactus* as *cac-tus*, and *whiskey* as *whis-key*. In addition, just as the lack of a definition of life (or death) does not prevent biologists from studying living things, the lack of a definition of the syllable should not prevent us from studying syllables. I shall argue that many questions about the syllable can be addressed, and reasonably answered, such as what the maximal syllable size is, what a possible onset is, and how to determine syllable boundaries. Such results constitute concrete progress towards the understanding of the syllable.

### 3.2. Maximal syllable size and word-edge consonants

Theories of syllables structure often assume a maximal syllable size for a given language.

Syllables within this size are in principle good. For example, if the maximal size is CCVCC, then CVCC, CCVC, CVC, and CV are generally good.

It is well known that word-medial syllables are generally quite simple, but extra consonants can occur at word edges. This is the case, for example, in Greek (Steriade 1982), English (Borowsky 1986), German (Giegerich 1985, 1989), Bella Coola (Bagemihl 1991), Spokane Salish (Bates and Carlson 1992), Polish (Bethin 1992), Georgian (Butskhrikidze 2002), and Jiarong (Lin 1993). Therefore, the maximal syllable size is mainly related to how we treat word-edge consonants. For example, consider word-final consonants in English. Several approaches are shown in (1).

#### (1) Rhyme size and the treatment of word-edge consonants

Analysis	<i>texts</i>	<i>helped</i>	Maximal rhyme
All-in	[teksts]	[helpt]	VCCCC
Suffix-out	[tekst]s	[help]t	VCCC
Coronals-out	[tek]sts	[help]t	VCC
Medial-based	[tek]sts	[hel]pt	VC
Two-step	[[tek]sts]	[[hel]pt]	VCCCC

The ‘all-in’ analysis is assumed by Jones (1950), Abercrombie (1967), Haugen (1956a, b), Fudge (1969), Hoard (1971), Kahn (1976), Hammond (1999), Hall (2002a), and Blevins (2003),

and many others. However, it has two problems. First, it must explain why medial syllables remain small. Second, it must explain why all sounds must be in a syllable. One might think that it is not possible to pronounce a consonant without a syllable, but English does have consonant interjections, such as *shh* [ʃ], *pff* [pf], and *psst* [ps], and English speakers have no trouble saying [s] or [f] alone. If such utterances are not syllables, then consonants can be pronounced without being in a syllable, and there is no need for the all-in analysis. If such utterances are syllables, there is another problem: a word like *texts* need not be one syllable but can be three [tek][s][ts], because [s] can be a syllable by itself (or along with [t]).

The ‘suffix-out’ analysis is assumed by Selkirk (1982), which yields smaller syllables than the all-in analysis. The ‘coronals-out’ analysis is assumed by Kiparsky (1981), among others. It is based on the observation that most word-final consonants in English are coronals [s, z, t, d, θ]. If we exclude them, the maximal rhyme is VCC. The ‘medial-based’ analysis is proposed by Giegerich (1985) and Borowsky (1986). It is based on the observation that medial rhymes are limited to VX (VV or VC). Therefore, more word-edge consonants are excluded.

While the suffix-out, coronals-out, and medial-based analyses yield a smaller syllable size, they face a common question, which is how to account for extra consonants at word edges. I shall argue below that they can be accounted for by morphology. If so, there is no need to assume a stretched syllable size at word edges. Instead, we can maintain a consistent syllable size for both edge and non-edge positions.

The ‘two-step’ analysis has been proposed by Kiparsky (1981), Steriade (1982), Borowsky (1989), and Giegerich (1992). It assumes that syllables are built in two steps. In the first, only core syllables are built and some word-edge consonants are excluded. In the second, word-edge consonants are absorbed by an adjacent syllable. Thus, Kiparsky’s analysis of *texts* is

[[tɛk]sts], where the inner brackets indicate the core syllable and the outer brackets indicate the final syllable. However, it is not obvious what the advantage of the two-step analysis is. For example, it does not explain why certain consonants can be absorbed but others cannot. In addition, if we can account for word-edge consonants by morphology, there is no need to assume that they have to be absorbed by or appended to a syllable. Moreover, if a syllable can take in more consonants, why does it not do so word medially? By the same argument, there is no advantage in assuming that word-edge consonants are appended to the word (Vaux 2004).

In summary, if we can account for word-edge consonants independently, we can maintain a consistent maximal syllable size, similar to that in word-medial positions.

### 3.3. Empty elements and the CV-only analysis

The maximal syllable size is also related to whether we assume empty elements. There are two kinds of empty elements in phonology. The first has phonetic realization. A well-known example is the pause at major phrase boundaries. Another example is the empty beat between stressed syllables, or at the end of a verse line. An example is shown in (2), where an empty beat is indicated by Ø.

(2) Empty elements that have phonetic realization (pauses)

*(Ding Ø) (dong Ø) (bell Ø) (Ø Ø)*

*(Kitty's) (in the) (well Ø) (Ø Ø)*

The rhythm in (2) is common in children's verse cross linguistically (Burling 1966). If one taps the lines, each line has four feet and each foot has two beats. The second beat of the first two feet in line 1 is empty, in that it does not correspond to a written syllable. Similarly, the last three

beats of each line are also empty. While empty beats are not indicated in orthography, they have phonetic realization, either as a pause or as the lengthening of the preceding syllable.

The second kind of empty elements have no phonetic content. For example, Lowenstamm (1996) and Scheer (2004) propose that both C and V can be empty, anywhere in a word. In addition, all sounds are syllabified into CV syllables. Two examples are shown in (3), where Ø is an empty C or V. In the CV-only analysis, most words that used to be called monosyllabic are now polysyllabic.

(3) The CV-only analysis

*mix*        [mɪ][kØ][sØ]

*spiked*    [sØ][pa][Øi][kØ][tØ]

The empty elements are purely abstract because there is in principle no contrast between [spa] and [sØ][pa], between [pai] and [pa][Øi], or between [kt] and [kØ][tØ]. The CV-only analysis necessarily requires [spa] to be represented as [sØ][pa], [pai] as [pa][Øi], and [kt] as [kØ][tØ]. In this regard, the use of Ø is theory internal (or circular).

The CV-only analysis has two problems. First, it must explain why only some word-edge Cs can occur with an empty V but others cannot. For example, why can [sØ] occur word initially in English but [fØ] cannot (e.g. *stop* is found but *\*flop* is not)? Similarly, it must explain why extra word-final Cs are mostly [s, z, t, d, θ] in English (e.g. *text* is good but *\*texp* is not). The CV-only analysis must say, as other theories do, that [s] is special initially and [s, z, t, d, θ] are special word finally. But once we agree that these sounds are special, it becomes unnecessary to assume that they must also occur with an empty V, or be syllabified at all.

Another problem with the CV-only theory is that it opens up alternative solutions that are hard to choose from. For example, consider three analyses of *spiked*, shown in (4).

(4) Possible analyses of *spiked* with purely empty elements

CVX-only     [sØØ][pai][kØt]

CV-only     [sØ][pa][Øi][kØ][tØ]

VC-only     [Øs][Øp][aØ][ik][Øt]

The CVX-only solution uses a larger syllable size but fewer syllables and fewer empty elements.

The CV-only and VC-only solutions use the same syllable size, the same number of empty elements, and the same number of syllables. All the solutions can represent any word without problem. It is not obvious how one should choose among the alternatives.

To conclude, there is no obvious advantage in using empty elements that do not have phonetic realization, nor is there any advantage of the CV-only theory.

### 3.4. The CVX theory

The CVX theory proposes that a word has the schematic structure in (5).

(5) The CVX analysis of word structure: C<sub>m</sub>CS...SCC<sub>m</sub>

S...S     one or more syllables, whose maximal size is CVX

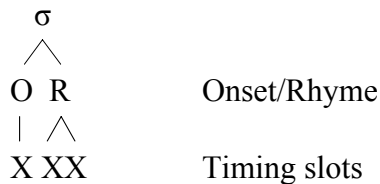
C     an optional C supported by a potential V (and anti-allomorphy)

C<sub>m</sub>     one (or more) optional affix or affix-like consonant

Let us begin with non-edge syllables. The CVX theory claims that in word-medial positions the maximal syllable contains three sounds, which can be CVV, such as [hau] *how* and

[bi:] *bee*, or CVC, such as [bæt] *bet*. I further assume that VX forms a constituent, which can be called the rhyme (although this constituent has been questioned by Davis 1988). There are three reasons for the constituent. First, it is well known that, while there are restrictions within the onset or within the rhyme, there are fewer restrictions between onsets and rhymes (Kessler and Treiman 1997). Second, the onset usually does not affect stress, whereas the coda often does (see Goedemans 1998, who argues that the onset never affects stress). Third, VX is the unit in syllable rhyming, so that *buy* [bai] rhymes with *why* [wai], to the exclusion of the onsets [b] and [w]. In contrast, there is no evidence that CV forms a unit to the exclusion of the coda in any phonological process. The structure of CVX is therefore given in (6).

(6) The maximal syllable: CVX



CVX is the shorthand for the structure in (6). The full structure itself does not indicate that the second slots must be filled by a vowel. Therefore, it is possible that the second slot is filled by a syllabic consonant. Some examples are shown in (7).

(7) Sample syllables

Transcription	C	V	X
[bæt]	[b	æ	t]
	<i>bat</i>		
[bai]	[b	a	i]
	<i>buy</i>		
[wai]	[u	a	i]
	<i>why</i>		
[jæm]	[i	æ	m]
	<i>yam</i>		

[bi:]	[b i i]	<i>bee</i>
[æn]	[ æ n]	<i>Ann</i>
[ŋ:]	[ ŋ ŋ]	‘fish’ (Shanghai Chinese)
[sz:]	[s z z]	‘four’ (Standard Chinese)

There are several points of interest. First, the C slot can be filled by a consonant, as in *bat*, or a high vowel (transcribed as a glide), as in *why* and *yam*. Second, the V slot can be filled with a syllabic consonant, as in [ŋ:] ‘fish’ in Shanghai Chinese and [sz:] ‘four’ in Standard Chinese. Third, the onset slot need not be filled, as in [æn] *Ann* and [ŋ:] ‘fish’ in Shanghai Chinese; I shall return to this issue below. Finally, [ii], [ŋŋ], and [zz] are not repetitions of the same sound but a sound doubly linked to two timing slots, represented in (8). A doubly linked sound has one articulatory configuration that lasts for two units of time.

(8)	CVX	VX	CVX
	∨	∨	∨
	b i	ŋ	s z
	[bi:]	[ŋ:]	[sz:]

In what follows I discuss the distribution of sounds inside CVX and sound strings that seem to exceed CVX.

### 3.4.1. Sonority, the peak rule, and the C rule

In this section I discuss the distribution of sounds within CVX. For example, given the sounds [ni], can they form a syllable [nni], where [n] is doubly linked to CV and [i] is linked to X?

Following a long tradition, I assume that sonority, or the loudness of a sound, governs the



distribution of sounds inside a syllable. According to Jespersen (1904), speech sounds can be ranked along the sonority scale in (9), where ‘>’ means ‘has greater sonority than’.

(9) Sonority scale (Jespersen 1904: 192):

low vowels > mid vowels > high vowels > r-sounds > laterals > nasals > voiced fricatives  
> voiced stops > voiceless fricatives > voiceless stops

It is possible to derive the sonority scale from phonological features. For example, among vowels, [+low] is more sonorous than [-low] and [-high] is more sonorous than [+high]. However, I shall omit the derivation (see Hooper 1976a, Kiparsky 1981, Steriade 1982, and Rice 1992 for such proposals). Within a syllable, the most sonorous sound should occur in the nucleus. Let us call it the peak rule, stated in (10).

(10) The peak rule:

The most sonorous sound should fill the V slot of CVX.

The most sonorous sound can also fill C or X, as long as it fills V. Some examples are shown in (11), where Ø indicates an unfilled onset.

(11) Illustration of the peak rule:

Input	CVX analysis	Peak rule violation
[bi]	[bii], *[bbi]	[i] > [b]
[sz]	[szz], *[ssz]	[z] > [s]
[ia]	[iaa], *[Øia], *[iia]	[a] > [i]
[i]	[Øii], [iii]	
[ŋ]	[Øŋŋ], [ŋŋŋ]	

[æŋ]            [Øæŋ], \*[ææŋ]            ??

In the first three cases, the bad forms violate the peak rule, because the sound in V is less sonorous than the sound X. For example, in \*[ssz], V is filled by [s], which is less sonorous than [z]. Also, the most sonorous sound can extend to C, as in [iii], and to X, as in [iaa] and [iii]. Finally, some input forms allow two (or more) analyses. For example, the input [i] can be syllabified as [Øii] or [iii]. In fact, [i] can also be syllabified as [Øi], where X is left out; the same is true for [ŋ]. However, \*[ææŋ] cannot be ruled out by the peak rule. If \*[ææŋ] is bad, we need a separate rule, which I call the C rule, stated in (12).

(12) The C rule:

The C slot of CVX cannot be filled by a non-high vowel.

Vennemann (1988: 70) makes an interesting observation that [rn] can be syllabified in two ways: either [r] is the nucleus, as in *lantern* [læŋ.tɹŋ], or [n] is the nucleus, as in *apron* [ei.pɹŋ]. If [r] has greater sonority than [n], the peak rule should rule out [pɹŋ]. In the CVX analysis, there is an explanation. Since there is only one onset slot, [pɹŋ] must be [p<sup>r</sup>ŋ], where [p<sup>r</sup>] is a complex sound. If the sonority of a complex sound is the same as the articulator that has greater closure (that is, the sonority of [p<sup>r</sup>] is the same as that of [p]), then [p<sup>r</sup>ŋ] no longer violates the peak rule.

The peak rule and the C rule suffice to determine the distribution of sounds inside CVX. In contrast, in other analyses, sonority is also used to analyze onset clusters, which I discuss next.

### 3.4.2. Onset clusters: sonority analysis vs. complex-sound analysis

Many languages allow onsets that are often transcribed as an obstruent-approximant cluster, such as [pr, pl, br, bl, fr, fl, kj, kw, ...]. Most analyses assume that onset clusters are governed by a sonority requirement (e.g. Selkirk 1982, Steriade 1982, and others), which I rephrase in (13), illustrated with English examples in (14).

(13) The sonority analysis of onset clusters:

The sonority in an onset cluster must show a sufficient rise.

(14) Examples of the sonority analysis of onset clusters

Cluster	Comment
pr	good: enough sonority rise
fr	good: enough sonority rise
θr	good: enough sonority rise
sr	good: enough sonority rise
*fn	bad: not enough sonority rise

The degree of sonority rise is measured by the sonority scale, such that the farther apart two sounds are on the scale, the greater sonority rise there is between them. More discussion of the sonority analysis of English onsets is given in Chapter 8.

In the CVX theory, there are no onset clusters. What appears to be an onset cluster is a complex sound (see Chapter 2). The proposal is given in (15) and exemplified in (16).

(15) The complex-sound analysis of onset ‘clusters’:

Two sounds can fit into the C slot of CVX only if they can form a complex sound.

- (16) a. Clusters that can form a complex sound: pr, fr, pl, kr, ...
- b. Clusters that cannot form a complex sound: θr, ʃr, ʃm, ʃn, sr, sl, sm, sn, st, sp, sk, sf, ...

In the sonority analysis, [θr] and [sr] are both good, because they have the same sonority rise as [fr]. In contrast, \*[θr] and \*[sr] are bad in the complex-sound analysis, even though [fr] is good. The reason is that [f] and [r] use different articulators and so can form a complex sound, whereas [θ] and [r] use different feature values of the same articulator, which violates the No Contour Principle (Chapter 2) and so they cannot form a complex sound. The feature structures of [fr] and [θr] are shown in (17) and (18).

- (17) [f]                      Labial—[+fricative]  
       [r]                      Coronal—[-fricative]  
       [f] + [r]              Labial—[+fricative], Coronal—[-fricative]
- (18) [θ]                      Coronal—[+fricative, +anterior]  
       [r]                      Coronal—[-fricative, -anterior]  
       [θ] + [r]              \*Coronal—[+fricative, -fricative, +anterior, -anterior]

For the same reason, [sr] (\*Coronal-[+fricative, -fricative]) and [fn] (\*Soft-palate-[-nasal, +nasal]) cannot form a complex sound or serve as an onset, nor can other clusters in (16b).

Since the sonority analysis and the complex-sound analysis make different predictions, we can test their claims. In particular, the sonority analysis predicts that clusters like [θr, ʃr] are possible word-medial onsets, because they have a proper sonority rise. In addition, clusters like

[ʃm, ʃn, sr, sl, sm, sn, st, sp, sk, sf] are possible word-medial onsets, because [s] (and [ʃ]) is exempt from the sonority requirement. In contrast, the complex-sound analysis predicts that such clusters are not possible onsets and will not be found word medially. We shall see in Chapter 8 that evidence supports the complex-sound analysis.

### 3.4.3. VVN and VNC rhymes

Although most non-final rhymes in English are limited to VX, sometimes VXC rhymes have been found (Borowsky 1989). Two well-known cases are VNC, where C is voiceless, such as [ɪmp] in *symptom*, and VVN, such as [aun] in *council*. In both cases the N has the same place of articulation as the following C. For example, in [mp] both sounds are Labial and in [ns] both sounds are Coronal.

Borowsky (1989) and Hall (2001) propose that an NC cluster with the same place of articulation can count as one sound. However, in feature theory NC clusters like [mp] and [ns] cannot be represented as a complex sound, because they contain the ill-formed structure Soft-palate-[+nasal, -nasal], which violates the No Contour Principle (Chapter 2). Therefore, Borowsky (1989) and Hall (2001) have not explained the VNC and VVNC problem. Goldsmith (1990: 151) proposes that English allows an extra sonorant between the nucleus and the coda, but the reason again remains unclear.

While NC cannot form a complex sound, VN can form a complex sound  $\tilde{V}$ , and so VNC becomes  $\tilde{V}C$  and VVNC becomes  $\tilde{V}\tilde{V}C$ . This analysis has three merits. First, the representation of VNV as  $\tilde{V}C$  is compatible with feature theory (Chapter 2). Second, the analysis agrees with a well-known fact that English vowels are nasalized before a nasal coda. Third, the analysis agrees with independent phonetic judgments that VNC is often realized as  $\tilde{V}C$ , such as *simple* [sɪp̃t̚],

*sinker* [sɪkə], *symptom* [sɪptəm], and *council* [kãüsɿ] (e.g. Malécot 1960, Bailey 1978, Fujimura 1979, and Cohn 1993). I return to this issue in Chapter 8.

#### 3.4.4. V:C rhymes

Another kind of VXC rhymes that appear in medial positions is V:C, such as [i:s] in *aesthetic* [i:sθetik] and [p:l] in *also* [p:lso] in American English. The vowel in V:C is tense, which can count as two sounds phonologically. For example, the rhyme of a stressed monosyllable must be either VC (as in *bit* [bɪt] and *bet* [bet]) or V: (as in *bee* [bi:] and *law* [lɔ:]), but not just V (e.g. \*[bɪ] or \*[bɛ]). This shows that a tense vowel is equivalent to a lax vowel plus a consonant.

If we treat a V:C rhyme as having three sounds and expand the rhyme size to VXC, we face two problems. First, there is a general lack of medial VCC rhymes. Second, there is a general lack of medial VVC rhymes, where VV is a diphthong (see Chapter 8). Therefore, V:C is a special case, not a general one.

Rather than stretching the rhyme size, a better solution is to assume that while tense vowels can be long, they need not in all environments (Pike 1947, Jones 1950, Abercrombie 1967, Giegerich 1985, and Alcantara 1998). On this view, V:C can be represented as VC, where V is tense and short but still distinct from a lax vowel. For example, *aesthetic* is [isθetik], where [i] is short (compared with [i:] in *bee*) yet distinct from [ɪ]. Similarly, *also* is [ɒlso], where [ɒ] is short compared with [ɒ:] in *author*. If this analysis is correct, the maximal rhyme is still VX.

## 3.4.5. Morphology and of word-edge consonants

Let us now turn to extra consonants at word edges. First, consider the C beyond a final VX, such as [p] in [help] *help* and [haip] *hype*. There are three proposals, shown in (19), where square brackets indicate syllable boundaries.

## (19) Analyses of final VXC

[...VXC]	Large rhyme
[...VX]<C>	Extrasyllabic
[...VX][C(V)]	Onset (of a potential vowel)

In the first analysis the entire VXC is in a large rhyme (e.g. Kahn 1976, Kiparsky 1981, Giegerich 1992, Harris 1994, Blevins 1995, and Hall 2001). The problem is that nonfinal rhymes are limited to VX (Giegerich 1985, Borowsky 1989), instead of VXC. Hall (2001) proposes that the rhyme is VX non-finally but VXC finally, but the proposal is a restatement of the fact, not an explanation. In addition, should there be VXC rhymes, one would expect them to be more prominent than VX rhymes, but there is no such evidence.

In the second analysis a word-final C is extrasyllabic (e.g. McCarthy 1979b, Hayes 1982, Borowsky 1989, Giegerich 1989, Goldsmith 1990, and Gussmann 2002). As a result, we obtain a consistent maximal rhyme VX for both final and nonfinal syllables. The question though is why the final C should be excluded from syllabification and why it is allowed to stay.

In the third analysis the final C is the onset of a following V. There are two views of what the V is. In the first, the V can be completely abstract (Burzio 1994). In the second, the V can be real—it is provided by a V-initial suffix (Giegerich 1985, Borowsky 1986). Since a word does not always have a suffix, I shall refer to the V as a ‘potential V’. For example, the [p] in *help* is

an extra C when the word occurs alone but it is the onset of the following V in *helper* and *helping*. The question for this analysis is why the final C is kept when there is no V-initial suffix, such as the [p] in *help*, *helpful*, and *helpless*. Here I appeal to a requirement known as paradigm uniformity or ‘anti-allomorphy’ (Burzio 1996), according to which one aims to keep a morpheme in the same shape regardless of the environment. In (20) I state the conditions that support the final C, which are illustrated in (21).

(20) Conditions for the final C

- a. Potential V: A word-final C can become the onset of a potential V, which comes with a V-initial suffix.
- b. Anti-allomorphy: Keep a morpheme in the same shape regardless of the environment.

(21) Final C	Supported by potential V	Supported by anti-allomorphy
[hɛl]p	[hɛl][pɪŋ], [hɛl][pə]	[hɛl]p, [hɛl]p[ʃəl], [hɛl]p[ɪs]
<i>help</i>	<i>helping, helper</i>	<i>help, helpful, helpless</i>
[rɪs]k	[rɪs][kɪŋ], [rɪs][ki]	[rɪs]k, [rɪs]k[fri]
<i>risk</i>	<i>risking, risky</i>	<i>rick, risk-free</i>

The analysis correctly predicts that a final C may behave as if it is extrasyllabic, in the sense that it is allowed even if the preceding rhyme is full, and that a final VC may behave like a final V (a light rhyme). In addition, since we can explain the existence of the extra C, there is no need to assume that it needs to join the preceding syllable. As a result, we can maintain a consistent maximal rhyme of VX for both final and nonfinal syllables.



A reviewer points out that in a word like *file*, the [l] is velarized (a ‘dark’ [l]), which is an indication that it is in the rhyme. Therefore, the syllabification should be [fail]. In the CVX analysis, [ail] is not a possible rhyme; instead, there are two other solutions. First, while it is true that [l] is ‘dark’ in the rhyme (for all speakers) and ‘clear’ in the onset (for some speakers), it is unclear whether [l] is dark or clear when it is unsyllabified. Second, a more likely answer is that *file* has two syllables [fai][ɫ], where [ɫ] is syllabic, in the rhyme, and hence dark.

The potential-V analysis also predicts that, if a language has CV prefixes, an extra C may occur as a ‘potential coda’ in word-initial position. This is schematically shown in (22).

(22) Initial C supported by potential V and anti-allomorphy

Root	Supported by potential V	Supported by anti-allomorphy
CCVC	[CV-C][CVC]	C[CVC]

I shall discuss in Chapter 11 that the Tibeto-Burman language Jiarong is such a case. Consider the examples in (23), from Lin (1993: 36), where [tɕ<sup>h</sup>] is an affricate.

(23) Root	Supported by potential V	Supported by anti-allomorphy
ntɕ <sup>h</sup> ok	[kɤ-n][tɕ <sup>h</sup> ok]	
‘dip’	‘dip’	
ʒba	[tə-ʒ][ba]	ʒ[ba-n][tɕ <sup>h</sup> ok]
‘face’	‘face’	‘face dip (dimple)’

The root for ‘dip’ is CCVC. When it follows a CV prefix (or a vowel-final word), the root-initial [n] can serve as the coda of the preceding V. Similarly, the root for ‘face’ is CCV. When there is

a CV prefix, the root-initial C can serve as the coda of the preceding V. When there is no prefix, as in ‘face dip (dimple)’, the root-initial C is supported by anti-allomorphy.

The potential-V analysis makes specific predictions for whether and where an extra C may occur in a given language. In a language that has V-initial suffixes, an extra C may occur in root-final position. In a language that has V-final prefixes, an extra C may occur in root-initial position. This is true for the languages we examine in the following chapters.

Given a potential V and anti-allomorphy, the maximal final string should be VXC. However, English words can take additional consonants beyond VXC. Some examples are shown in (24).

(24) Final consonants beyond VXC

Word	Sounds	Beyond VXC
<i>saved</i>	[seivd]	[d]
<i>risked</i>	[riskt]	[t]
<i>text</i>	[tekst]	[t]
<i>texts</i>	[teksts]	[ts]
<i>sixths</i>	[sɪksθs]	[θs]

It has been observed that final consonants beyond VXC are limited to [t, d, s, z, θ] in English. Phonetically, such sounds are all made with the Coronal articulator. One might suggest, therefore, that Coronal sounds are special. Let us call it the coronal hypothesis and state it in (25).

## (25) The coronal hypothesis:

Coronal obstruents do not need to be supported by a syllable, or they can be exceptionally attached to a syllable, even if the syllable is full.

On the other hand, it can be observed that [t, d, s, z, θ] are used as consonant suffixes in English; indeed they are the only consonant suffixes in English. This fact suggests a different explanation for the final consonants: they can survive not because they are in a syllable but because they are suffixes. A similar point is made by Goldsmith (1990: 127), who proposes that consonant morphemes are ‘licensed’ to occur even if they are not in a syllable. I shall call it the affix rule and state it in (26).

## (26) The affix rule (preliminary version):

Affix sounds can be pronounced, regardless of whether they fit into a syllable.

The affix rule is intuitively natural. According to the CELEX lexicon of English (Baayen et al 1993), among the 54,447 basic words, 41,911 end in C. This means that in most English words the final syllable is full. If consonant suffixes are not pronounced when the preceding syllable is full, then most of the time we would not be able to tell whether there is a suffix or not, and the use of suffixes would become quite irrelevant.

Whether one assumes the coronal hypothesis or the affix rule, there is no need to expand the rhyme size VX, because the sounds are already accounted for. Therefore, it is redundant to assume that they must also be absorbed into or hosted by a syllable.

The coronal hypothesis and the affix rule make different predictions. According to the former, we should find extra coronal consonants not only word finally but also word medially. In

contrast, according to the affix rule, we should find extra coronals only at word edges. Evidence from English, German, and Jiarong supports the affix rule, to be seen in Chapters 8, 10, and 11.

In English, not all final consonants beyond VXC are suffixes; for example, the final [t] in *text* [tekst] is not a suffix. Similarly, the final [st] in the German word *Herbst* [ɛrpst] ‘autumn’ are not suffixes. The coronal hypothesis could say that such extra consonants are allowed because they are coronals. If we assume the affix rule, what would the explanation be?

One proposal is that, since both German and English use coronal consonants as suffixes, all final coronals are perceived as suffixes (Fujimura 1979, Pierrehumbert 1994). Another way to look at the matter is that, because we are used to hearing coronal suffixes, other final coronals would not sound bad to the ear. The two proposals are stated in (27) and (28).

(27) Perceived affixes:

Sounds that resemble affixes can be treated as affixes.

(28) Phonetic familiarity:

If morphology requires one or more sounds  $C_m$  to occur in a given phonological environment (e.g. word-final position beyond VXC), then we accept  $C_m$  in the same phonological environment even if the morphological requirement is absent.

If either proposal is correct, we can revise the affix rule as in (29).

(29) The affix rule (final version):

Affix or affix-like sounds can be pronounced, whether they can fit into a syllable or not.

I shall show in Chapter 8 and Chapter 10 that many words in English and German can be accounted for by the affix rule.

## 3.4.6. Can [lp] be an onset or [pl] be a coda?

I suggested above that a word like *plum* is not CCVC but CVC, because [p] and [l] can form a complex sound [p<sup>l</sup>] (Chapter 2). Several questions can be raised. First, why is there no word *lpum*? The word ought to fit into CVC, if [l] and [p] can form a complex sound. The answer is that all articulatory features in a complex sound are simultaneous, so that [p]+[l] and [l]+[p] are identical. Therefore, *lpum* would be the same as *plum*.

One might also ask why *help* is not realized as [hɛp<sup>l</sup>], where [l] and [p] form a complex sound [p<sup>l</sup>] in the coda. The answer is that [p] is already supported by ‘anti-allomorphy’ (see above), and so there is no need to incorporate it into a syllable. Besides, the sequence [lp] is more similar to that in *hel.per*, and hence is a better way to satisfy anti-allomorphy.

Next, one might ask why there are no such words as *tikl* or *tepl* in English, where [k<sup>l</sup>] and [p<sup>l</sup>] are complex sounds in the coda. The answer is that [l] can be syllabic, and *tikl* and *tepl* would be disyllabic. Indeed, there are such words in English, such as *tickle* and *nipple*. This point was made by Lamontagne (1993).

Finally, one might ask why there is no medial coda [p<sup>l</sup>] or [k<sup>l</sup>], which could result from a hypothetical word *neplsa* or *tikl<sup>l</sup>ny*. A possible answer is that, because [l] can be syllabic, *neplsa* or *tikl<sup>l</sup>ny* can be trisyllabic (likely to be spelled as *neppelsa* or *tickelny*). Another possible answer is that languages generally use fewer codas than onsets. For example, Cantonese uses about fifteen consonants in the onset but six in the coda: [p, t, k, m, n, ŋ]. Similarly, Standard Chinese allows about twenty consonants in the onset but two in the coda: [n, ŋ]. Therefore, it is possible that English simply allows [p<sup>l</sup>] and [k<sup>l</sup>] in the onset but not in the coda.

### 3.4.7. Summary

There is little doubt that in many languages the maximal syllable size is at least CVX. What I have proposed is that CVX is also the upper limit on syllable size, where C, V, and X can each be a complex sound. Extra consonants can be found at word edges, but they can be explained by morphology. In particular, a word has the schematic structure  $C_mCSCC_m$ , where  $C_m$  is one or more affix or affix-like consonants, C is a consonant supported by a potential V from an affix, and S is one or more syllables whose maximal size is CVX.

English has many V-initial suffixes and a small number of V-final prefixes. In addition, English has consonant suffixes but not consonant prefixes. Therefore, the maximal English monosyllable is shown in (30).

#### (30) Maximal monosyllable in English

$CSCC_m$

S: maximally CVX, where C, V, and X can be a complex sound

A similar analysis applies to German. Comprehensive quantitative data from English and German will be provided in Chapter 8 and Chapter 10 in support of the analysis.

As in previous analyses, the CVX theory predicts cross-linguistic variations in the maximal size of a monosyllabic word. However, unlike previous analysis, which attributes the variations to syllable parameters, the CVX theory derives the variations from morphology. Let us call the CVX theory the ‘morphological’ approach and the traditional analysis that assumes syllable parameters the ‘phonological’ approach. In (31) I compare the two, where the ‘phonological’ approach is similar to that of Blevins (1995).

## (31) Accounting for variations in the size of monosyllables (S is maximally CVX)

Max size	Morphological (CXV)	Phonological (parameters)
S	No V-final prefixes, V-initial suffixes, or C affixes	No branching onset or branching coda
CS	Having V-final prefixes	Branching onset
C <sub>m</sub> S	Having C prefixes	Branching onset; appendix
SC	Having V-initial suffixes	Branching coda
SC <sub>m</sub>	Having C suffixes	Branching coda; appendix
CSC	Having V-final prefixes and V-initial suffixes	Branching onset and branching coda
SCC <sub>m</sub>	Having V-initial suffixes and C suffixes	Branching coda and appendix
Etc.		

The main claim of the CVX theory is that there is a correlation between the morphology of a language and the maximal size of monosyllabic words. The correlation has not been proposed before, but I shall show that it is true for the languages discussed in the following chapters. If the correlation is true in general, then the ‘phonological’ approach becomes redundant, because the morphological properties must be recognized anyway. In (32) I summarize the differences between the two approaches.

(32)	CVX	Traditional (syllable parameters)
Predicting variation	yes	yes
Accounting for extra Cs	yes	yes
Morphological variation	yes	(yes)
Syllable parameters	no	yes
Consistent syllable size	yes	no
Over-prediction	no	yes

Both analyses account for the cross-linguistic variation in the size of the monosyllable and for extra consonants at word edges. However, the CVX analysis assumes no syllable parameters and achieves a more consistent syllable size. In contrast, since traditional analyses must assume morphology variation anyway, their assumption of syllable parameters becomes redundant. In addition, traditional analyses assume a wider range of syllable sizes and a much greater amount of over-prediction.

### 3.5. Syllable boundaries

If a language allows CVC, CV, and VC syllables, as English does, how would CVCVC be syllabified? Both CV.CVC and CVC.VC seem to be possible. But if a given word is syllabified in just one way, we need a way to resolve the ambiguity. I discuss several proposals.

#### 3.5.1. Speaker intuition

According to Chomsky (1957), a grammar is a set of rules that define possible and impossible linguistic structures. In addition, knowing the grammar, a native speaker has the intuition to judge whether a structure is or is not good in his or her language. Similarly, Haugen (1956a, b)



proposes that speakers can break up words into syllables if they are asked to speak slowly, and Halle (1962) argues that speakers have the intuition to judge whether a sound sequence is well formed, even if it is not a real word. For example, [bɪk], [θod], [gnait], and [vnig] are (or were) not real English words but the first two are possible words while the last two are not.

However, recent studies (e.g. Frisch et al 2000, Myers and Tsay 2005, and Zhang 2007) show that speaker judgment on possible words is not always clear cut. In addition, many studies have noted that consistent judgment on syllable boundaries can be hard to obtain (Gimson 1970, Treiman and Danis 1988, Giegerich 1992, Hammond 1999, Steriade 1999, Blevins 2003).

Giegerich (1992) proposes that speaker judgment can be probed with a better test. In the test speakers are asked to pronounce each syllable of a word twice. One can begin with words whose syllable boundaries are unambiguous, such as *after*, which is pronounced as *af-af-ter-ter*. More difficult words can then be presented to see how speakers deal with them. One such word is *apple*, whose output is reported to be *ap-ap-ple-ple*. Giegerich takes the result to mean that *apple* is syllabified as *ap-ple*, where the [p] is ambisyllabic.

There are several questions about this proposal. First, if the output for *apple* is *ap-ap-ple-ple*, instead of *ap-ap-le-le*, does it mean that the syllables are *ap-ple*, or does it mean that *le-le* [l-] is an unusual sequence of syllables which the speaker would avoid? Second, consider the word *text*. If the result is *tek-tek-st-st*, one might conclude that the syllable is [tɛk] and [st] is outside the syllable. However, if the result is *text-text*, does it mean that the syllable is [tɛkst], or does it mean that the speaker is trying to avoid repeating a non-syllable cluster [st]? Finally, Giegerich does not discuss whether speaker judgment is always clear. My own test with some native

speakers shows that the judgment can vary. For example, the output for *city* can be *cit-cit-ty-ty* or *ci-ci-ty-ty*. Therefore, the test does not seem to provide conclusive answers.

But suppose there is a total lack of speaker intuition, should we conclude that there are no syllable boundaries? I suggest the answer be no. The reason is that not everything is intuitively obvious. For example, we are not aware of how we see colors, how we digest food, or how we walk. Similarly, if we rely on intuition, we would wrongly conclude that the earth is flat.

### 3.5.2. The Law of Initials (LOI) and the Law of Finals (LOF)

Since a word often begins with a syllable, it is natural to expect word-medial onsets to resemble word-initial onsets. Similarly, since a word often ends with a syllable, it is natural to expect word-medial rhymes to resemble word-final rhymes. As Haugen (1956a: 196) put it, the idea can be used for ‘making syllabic cuts on the basis of initial and final clusters’. Two formulations of the idea are shown in (33) and (34). For familiarity, I use ‘syllable onsets’ for what Vennemann calls ‘syllable heads’.

(33) Vennemann (1988: 32-33):

- a. Law of Initials: Word-medial syllable heads are the more preferred, the less they differ from possible word-initial syllable heads of the language system.
- b. Law of Finals: Word-medial syllable codas are the more preferred, the less they differ from possible word-final syllable codas of the language system.

(34) Blevins (2003: 203):

- a. If  $C_0$  is possible word-initially, then  $C_0$  is possible syllable-initially.
- b. If  $C_0$  is not possible word-initially, then  $C_0$  is not possible syllable-initially.

- c. If  $C_0$  is possible word-finally, then  $C_0$  is possible syllable-finally.
- d. If  $C_0$  is not possible word-finally, then  $C_0$  is not possible syllable-finally.
- e. If  $V_q$  is possible word-initially, then  $V_q$  is possible syllable-initially.
- f. If  $V_q$  is not possible word-initially, then  $V_q$  is not possible syllable-initially.
- g. If  $V_q$  is possible word-finally, then  $V_q$  is possible syllable-finally.
- h. If  $V_q$  is not possible word-finally, then  $V_q$  is not possible syllable-finally.

According to Vennemann, word-medial onsets and rhymes should be similar to those at word edges, but not vice versa. For example, the word-final rhyme VCCC in *text* does not have to resemble a word-medial rhyme, presumably because no VCCC rhyme is found word-medially. In contrast, Blevins (2003) expects all word-initial consonant clusters to occur as word-medial onsets and all word-final consonant clusters to occur as word-medial codas. For example, because [-kst] occurs word-finally in *text*, Blevins expects it to be a possible coda word-medially, but no such medial codas are found. In this regard, Blevins makes a lot of over-predictions.

If we exclude extra consonants at word edges, we can maintain a two-way resemblance relation between word-edge syllables and those in word-medial positions. Therefore, I define the relation in (35).

- (35) The relation between word-edge syllables and those in word-medial positions:
- a. Law of Initials (LOI): Word-initial onsets and word-medial onsets should resemble each other.
  - b. Law of Finals (LOF): Word-final rhymes and word-medial rhymes should resemble each other.

The LOI and LOF can interact with other factors. For example, if a language does not have final stress, it may disallow word-final C, yet still have CVC.CV words, where the first syllable violates the LOF. However, if we leave aside such factors, we would expect the LOI and LOF to be true overall.

The new definition can help syllabification both word-medially and at word edges. For example, because medial rhymes are limited to VX, the rhyme in *text* should not include [st]. Similarly, because there is no word-initial onset [ft] or word-final [ε], we would syllabify *hefty* as *hef.ty*, not *he.fty*. Some additional examples are shown in (36).

(36) Syllabification by LOI and LOF

[sɪŋ.ə] <i>singer</i>	OK: <i>sing</i> , <i>Urbana</i> .
*[sɪ.ŋə] <i>singer</i>	Violating LOI: [ŋ-] is not a word-initial onset
	Violating LOF: [ɪ] is not found word-finally
[sɪt.i] <i>city</i>	OK: <i>sit</i> , <i>east</i> (note: word-final [t] can flap, as in <i>sit in</i> ).
*[sɪ.ti] <i>city</i>	Violating LOF: [ɪ] is not found word-finally
[æt.ləs] <i>atlas</i>	OK: <i>at</i> , <i>Las Vegas</i> .
*[æt.l.əs] <i>atlas</i>	Violating LOF: [-tl] is not a word-final coda.
*[æ.tl.əs] <i>atlas</i>	Violating LOI: [tl-] is not a word-initial onset.

\*[wɪ.ski] *whiskey* Violating LOF: [-ɪ] is not a word-final rhyme.

[wɪs.ki] *whiskey* OK: *miss, Kyoto*.

\*[wɪsk.i] *whiskey* OK: *risk, east*, but rhyme exceeds VX.

[taɪ.ni] *tiny* OK: *tie, need*.

\*[taɪn.i] *tiny* OK: *fine, east*, but rhyme exceeds VX.

The LOI and the LOF can unambiguously determine syllabification in *singer*, *city*, and *atlas*. In *whiskey* and *tiny*, there is ambiguity if the rhyme length is not limited. To resolve the ambiguity, Vennemann (1988) proposes some additional laws. Of interest are the Onset Law (Head Law), which favors a single-sound onset, the Nucleus Law, which favors a single-sound nucleus, and the Coda Law, which favors a single-sound coda. These laws in effect favor a CVX syllable and can resolve the ambiguity in *whiskey* and *tiny*.

However, the LOI and the LOF cannot resolve another kind of ambiguity, even when CVX is assumed. This is shown in (37).

(37) Ambiguity under LOI and LOF:

\*[kæ.nəd.ə] *Canada* Violating LOF: [-æ] is not a word-final rhyme.

[kæn.əd.ə] *Canada* OK: *can, admitted, a*.

[kæn.ə.də] *Canada* OK: *can, a, Dakota*.

To resolve the ambiguity in (37), additional assumptions are needed.

### 3.5.3. Maximizing stressed syllables

Hoard (1971) proposes that ‘stressed syllables attract a maximum number of segments’. Bailey (1978) also proposes that there should be ‘more consonants in a cluster to be syllabified with the heavier-accented of two surrounding nuclei as the tempo increases’. Similar proposals are made by Wells (1990) and Hammond (1999). Some examples are shown in (38).

#### (38) Maximizing stressed syllables

Bailey (1978): *tiger* [taɪg.ə], *capital* [kæp.ət.l], *ability* [ə.bɪl.ət.i], *multiply* [mʌlt.ə.plaɪ]

Wells (1990): *city* [sɪt.i], *dolphin* [dɒlf.ɪn], *cauldron* [kɔːldr.ən]

Hammond (1999): *fealty* [fɪlt.i], *bulky* [bʌlk.i], *alcove* [ælk.kov]

There are two problems with this approach. The first is that it predicts many large rhymes that exceed VX, but such clusters are generally lacking (see Chapter 8). Secondly, it yields ambiguity when syllabifying unstressed syllables. For example, *Canada* can be [kæn.əd.ə] or [kæn.ə.də]. In both cases the stressed syllable is maximized, but the theory does not say whether an unstressed syllable can or cannot take a coda.

### 3.5.4. Maximal Onset

The Maximal Onset rule was proposed by Pulgram (1970) and Kahn (1976), according to which consonants between vowels should be syllabified as the onset of the following vowel as far as possible. Under the Maximal Onset rule, there is no ambiguity in syllabification. Some examples are shown in (39).

## (39) Syllabification by Maximal Onset

[sɪŋ.ə] <i>singer</i>	*[sɪ.ŋə], because [ŋ-] is not a possible onset
[æt.ləs] <i>atlas</i>	*[æt.tləs], because [tl-] is not a possible onset
[taɪ.ni] <i>tiny</i>	*[taɪn.i], because [n-] is a possible onset
[lɛ.mən] <i>lemon</i>	*[lɛm.ən], because [m-] is a possible onset
[kæ.nə.də] <i>Canada</i>	*[kæn.əd.ə], because [n-] and [d-] are possible onsets

Steriade (1982: 77) and Clements and Keyser (1983: 37) further assume that VCV is syllabified as V.CV universally, a view that is shared by many others.

One problem with the Maximal Onset analysis is that it does not respect the Law of Finals. For example, in [lɛ.mən] *lemon* the first syllable ends in [ɛ], which is not found word-finally. To avoid this problem, Pulgram (1970) and Kahn (1976) added a condition that in VCV, if the first V is lax and stressed, the syllabification is VCV, where C belonging to both the first and the second syllables. Similarly, Selkirk (1982) proposes a rule to change V.CV to VC.V if the first V is lax and stressed. However, many other studies disregard the Law of Finals and use the Maximal Onset rule strictly (e.g. Halle and Vergnaud 1987 and Baayen et al 1993). A second problem with the Maximal Onset analysis is that there is no clear evidence for it.

For example, it is well-known that vowel-initial words often start phonetically with a glottal stop. Two examples in English are shown in (40).

- (40) *out*    [ʔaut]  
*Ann*    [ʔæn]

The fact is often interpreted as evidence that every syllable needs an onset, and when there is no consonant onset, a glottal stop is added.

However, there are two ways to interpret the initial glottal stop. One is that it is an intended sound. The other is that it is an unintended gesture: the vocal tract cannot assume the vowel gesture all of a sudden, and the glottal stop reflects an unintended state before the vowel is pronounced. According to the first interpretation, the glottal stop should be present whether there is a preceding word or not. According to the second, the glottal stop should be absent when there is a preceding word. Evidence supports the second view, as the example in (41) shows.

(41) No glottal stop

*ran out*            [rænaut] / \*[rænʔaut]

When a vowel-initial word follows another word, the glottal stop cannot be added (unless one is speaking very slowly). One might suggest that perhaps the final consonant of the first word has been shifted to the onset of the second word. However, English generally does not re-syllabify across words. For example, there is a distinction between each pair of expressions in (42), cited in Ladefoged (1972) and Wells (1990).

(42) *(I'm going to get my) lamb prepared* vs. *(I'm going to get my) lamp repaired*

*nitrate* vs. *night-rate*

*plum pie* vs. *plump eye*

Similarly, word pairs like those in (43) remain distinct (the first pair is from Kiparsky 1979). In the examples, vowel length is not shown and only relevant syllable boundaries are shown.



- (43) *at ease*            [æɾ.iz] / \*[æ.t<sup>h</sup>iz] / \*[æt.ʔiz]  
          *a tease*            [ə.t<sup>h</sup>iz]  
          *beef eater*        [bif.irə̃] / \*[bif.ʔirə̃]  
          *bee feeder*        [bi.firə̃]  
          *beat owls*        [bir.aulz] / \*[bi.t<sup>h</sup>aulz]  
          *be towels*        [bi.t<sup>h</sup>aulz]

In American English [t<sup>h</sup>] and [ɾ] are allophones of [t], where [t<sup>h</sup>] is used in the onset of a stressed syllable. In *at ease*, [t<sup>h</sup>] cannot be used, which means that [t] remains in the first syllable; [ʔ] cannot be added either, which means that the onset is not required. The difference between *beef-eater* and *bee-feeder* makes the same point, namely, there is no cross-word syllabification of [f] in the former. If there is, the two compounds should sound the same (the phonetic difference seems to lie in the length of the first [i], which is longer in *bee* than in *beef*). In addition, there is no [ʔ] in *beef-eater* either, which shows again that the onset is not required. The same analysis applies to the difference between *beat owls* and *be towels*.

In summary, there is no clear evidence for the Maximal Onset rule. Kahn (1976) argues that the Maximal Onset analysis is supported by other allophonic rules, such as aspiration and flapping. We will see below that the evidence is not compelling either.

## 3.5.5. The Weight-Stress Principle (WSP)

The WSP has been proposed in various forms by many people, such as Prokosch (1939), Fudge (1969), Hoard (1971), Bailey (1978), Selkirk (1982), Murray and Vennemann (1983), Kager (1989), Prince (1990), Wells (1990), Hammond (1999), and Duanmu (2000). There is also some phonetic evidence for the WSP (Krakow 1989, Turk 1994, and Redford and Randall 2005).

Counter examples to the WSP have been reported (e.g. Davis 1988: 85-88), but the evidence does not seem compelling and is not discussed here. The WSP is stated in (44) and illustrated in (45), where V is a short stressed vowel, X can be C or V, and v is an unstressed vowel.

(44) The Weight-Stress Principle (WSP):

- a. Stressed syllables should be heavy (rhyme being VX).
- b. Unstressed syllables should be light (rhyme being v or a syllabic C).

(45)	Good	Bad	Reason	Example (relevant part underlined)
	VC.v	*V.Cv	WSP-a	<i>ci<u>ty</u></i>
	VC.VX	*V.CVX	WSP-a	<i>ra<u>bbi</u></i>
	v.CVX	*vC.VX	WSP-b	<i><u>atta</u>ck</i>
	v.Cv	*vC.v	WSP-b	<i>Ca<u>na</u>da</i>
	VV.Cv	*VVC.v	Max VX	<i>co<u>la</u></i>
	VC.Cv	*VCC.v	Max VX	<i>whi<u>sper</u></i>

The WSP offers unambiguous syllabification for the first four cases. In the last two cases there is no ambiguity either, because the VX limit on rhyme size can exclude VVC.v and VCC.v. It can be seen, too, that the WSP analysis satisfies both the Law of Initials and the Law of Finals. Finally, we shall see below that the WSP analysis can account for allophonic variations, too.

The WSP raises a question: How much information is there in the lexicon? Some studies assume that the English lexicon contains no syllables or stress; instead, we use rules to build syllables and assign stress (e.g. Halle 1962, Halle and Vergnaud 1987, and Zec 1988). Other studies assume that the lexicon can contain a lot of memorized information (e.g. Vennemann 1974, Selkirk 1980, Clements and Keyser 1983, Giegerich 1992, Bromberger and Halle 1989, Burzio 1996, Hayes 1995, Halle 1997, Pater 2000, Ladefoged 2001, Pierrehumbert 2001, Vaux 2003, Port 2006, and Coleman and Pierrehumbert 2007). I follow the latter and assume that syllable structure or stress can be part of memorized information in the lexicon.

### 3.5.6. Aspiration, flapping, and sounds at word-medial syllable boundaries

Kahn (1976) argues that syllabification can help account for allophonic rules, such as aspiration and flapping in American English. His analysis is summarized in (46) and exemplified in (47), where C is ambisyllabic, which means that it belongs to both the first and the second syllable.

#### (46) Kahn's rules for syllabification, aspiration, and flapping

$VCV \rightarrow V.CV$       Maximal Onset rule

$V.Cv \rightarrow VC\underline{v}$       Ambisyllabic rule (v is unstressed vowel)

Aspiration: [p, t, k] are aspirated when they start a syllable

Flapping: Intervocalic [t] or [d] is flapped when ambisyllabic

#### (47) Analysis of *potato*

[pə.té:.to]      Maximal Onset rule

[pə.té:to]      Ambisyllabic rule

[p<sup>h</sup>ə.t<sup>h</sup>é:ro]      Aspiration and flapping

In *potato*, the final vowel [o] is unstressed, and so the ambisyllabic rule applies to the second [t], giving [pə.té.to]. The analysis correctly predicts aspiration and flapping. However, Kahn's analysis faces a problem with *af.ter*, where [t] starts the second syllable (because [ft] is not an onset in English), yet [t] is not aspirated. Kahn proposes that *after* can be syllabified as *a.ftər*, where [ft] is an onset, but there is little evidence for the claim.

Borowsky (1986) abandons the ambisyllabic rule and adopt a different resyllabification rule and a different interpretation for flapping. Her analysis is shown in (48) and (49).

(48) Borowsky's rules for syllabification, aspiration, and flapping

VCV → V.CV      Maximal Onset rule

V.Cv → VC.v      Resyllabification (v is unstressed vowel)

Flapping: Intervocalic [t] or [d] is flapped when it is in the coda

(49) Analysis of *potato*

[pə.té:.to]      Maximal Onset rule

[pə.té:t.o]      Resyllabification

[p<sup>h</sup>ə.t<sup>h</sup>é:r.o]      Flapping

Because [o] is unstressed, resyllabification moves the second [t] to the coda of the preceding vowel, giving [pə.té:t.o]. Flapping then applies to the second [t] but not the first, which is in the onset. It is worth noting that Borowsky's resyllabification rule can violate the VX rhyme size, which she otherwise assumes. For example, the second rhyme in *potato* is [e:r], which is VVC, because a tense vowel counts as VV. Similarly, the [t] in *mighty* is flapped, and so the

syllabification must be [maɪr.i], where the first rhyme again exceeds VX. The VVC rhyme has also been proposed by Hoard (1971), Selkirk (1982), Wells (1990), and Hammond (1999).

Jensen (2000) proposes yet another analysis of aspiration and flapping (see also Iverson and Salmons 1995, Davis 1999, Davis and Van Dam 2001), in which there is no need for ambisyllabic sounds or resyllabification. Instead, the Maximal Onset rule is sufficient. His analysis is shown in (50) and (51).

(50) Jensen's rules for syllabification, aspiration, and flapping

VCV → V.CV                      Maximal Onset rule

Aspiration: [p, t, k] are aspirated word initially or foot initially

Flapping: Intervocalic [t] or [d] is flapped when it is not word initial or foot initial

(51)    *potato*                      *at ease*

[pə.té:.to]                      [æt.íz]                      Syllabification by Maximal Onset

[p<sup>h</sup>ə.t<sup>h</sup>é:.ro]                      [æɾ.íz]                      Aspiration and flapping

In *potato*, [p] is aspirated because it is word initial. The first [t] is aspirated because it is foot initial, assuming that each stressed syllable starts a foot. The second [t] is flapped because it is neither word initial nor foot initial. Similarly, in *at ease*, the [t] is in the coda of the first vowel, assuming no resyllabification across words. The [t] is flapped because it is neither word initial nor foot initial. Thus, [t] can be flapped whether it is in the onset, as in *potato* [p<sup>h</sup>ə.t<sup>h</sup>é:.ro], or in the coda, as in *at ease* [æt.íz], and whether it precedes a stressed vowel, as in *at ease* [æt.ís], or an unstressed vowel, as in *hit it* [hír.ɪt]. Jensen has no problem with *after* [æf.tə] either, because

the [t] is not foot initial, so it is unaspirated.

As Kahn (1976), Borowsky (1986), and Jensen (2000) observe, aspiration and flapping are sensitive to stress. If we recognize the role of stress, we can derive syllabification without the Maximal Onset rule. The solution, shown in (52) and (53), is to use the Weight-Stress Principle (WSP) for syllabification, and the same rules for aspiration and flapping as Jensen (2000). A comparison between the WSP analysis and others is shown in (54).

(52) Syllabification by WSP, aspiration, and flapping

WSP:  $VCv \rightarrow VC.v$ ,  $vCV \rightarrow v.CV$

Aspiration: [p, t, k] are aspirated word initially or foot initially

Flapping: Intervocalic [t] or [d] is flapped when it is not word initial or foot initial

(53) *potato*                      *at ease*

[pə.té:.to]                      [æt.íz]                      Syllabification by WSP

[p<sup>h</sup>ə.t<sup>h</sup>é:.ro]                      [ær.íz]                      Aspiration and flapping

(54) Comparison of four analyses

<i>at ease</i>	<i>city</i>	<i>potato</i>	
[ær.íz]	[síri]	[p <sup>h</sup> ə.t <sup>h</sup> é:.ro]	Kahn (1976)
[ær.íz]	[síri]	[p <sup>h</sup> ə.t <sup>h</sup> é:r.o]	Borowsky (1986)
[ær.íz]	[sí.ri]	[p <sup>h</sup> ə.t <sup>h</sup> é:.ro]	Jensen (2000)
[ær.íz]	[síri]	[p <sup>h</sup> ə.t <sup>h</sup> é:.ro]	WSP

All the analyses agree that in *at ease* there is no resyllabification across words, but they differ in the analysis of *city* and *potato*. In *city*, [r] is ambisyllabic for Kahn, in the coda for Borowsky and the WSP analysis, and in the onset for Jensen. In *potato*, [r] is ambisyllabic for Kahn, in the coda for Borowsky, and in the onset for Jensen and the WSP analysis. In particular, in the WSP analysis, the first [t] in *potato* is not a coda, because the first syllable is unstressed and must be light. The second [t] is not a coda either, because [e:] is long and has filled up the VX rhyme. While the WSP analysis offers different syllabification, it accounts for aspiration and flapping just as well.

Hoard (1971) observes some other examples: *motto*, where the [t] is flapped, and *veto*, *Hittite*, and *satire*, where the medial [t] is not flapped. The WSP analysis is shown in (55).

(55)	<i>motto</i>	[ma:][ro] or [mar][o]
	<i>veto</i>	[vi:][t <sup>h</sup> o:]
	<i>Hittite</i>	[hit][t <sup>h</sup> ai](t)
	<i>satire</i>	[sæ:][t <sup>h</sup> ai][ʔ] or [sæt][t <sup>h</sup> ai][ʔ]

In *motto*, the second syllable is unstressed and so [t] is not foot-initial. In contrast, the second syllable in *veto*, *Hittite*, and *satire* is stressed and so its onset [t] is foot-initial and aspirated. The analysis echoes Hoard's proposal that sometimes a geminate consonant can occur in English, such as [tt] in *Hittite*.

Blevins (2003) argues that some words cannot satisfy the Law of Initials or the Law of Finals, no matter how we syllabify them. One example is *lemon*, shown in (56).

(56)	Syllabification	Problem
	[lɛ.mən]	Violating LOF: No word-final [ɛ]
	[lɛm.ən]	Violating LOI: No word-initial [ə]?

In [lɛ.mən] there is a syllable-final [ɛ], which is not found word-finally. In [lɛm.ən] there is a syllable-initial [ə], which Blevins claims is not found word-initially, because she assumes that word-initial vowels are preceded by a glottal stop. However, Blevins's claim is true only for words pronounced in isolation, but not for words pronounced with another word. For example, while *out* can start with a glottal stop in isolation, there is no glottal stop in *ran out* [rænaut]. Thus, if we consider words in context, then *lem.on* satisfies both the LOI and the LOF and is the only correct syllabification. Similarly, consider the analysis of *Betty*, shown in (57).

(57)	Syllabification	Word in isolation	Word in context
	[bɛr.i]	Final [r]; initial [i]	OK: <i>at ease</i>
	[bɛ.ri]	Final [ɛ]; initial [r]	Violating LOF: no word-final [ɛ]

If we look at words in isolation, both syllabifications are problematic, because [r] and [ɛ] are not found at word-finally, and [i] and [r] are not found at word-initially. However, if we look at words in context, then [bɛr.i] is unproblematic but [bɛ.ri] is. In [bɛr.i], the syllable-final [r] and the syllable initial [i] are found at word edges, as in *at ease*. In contrast, in [bɛ.ri], the syllable-final [ɛ] is not found word-finally.



There are, however, some cases where word-edge patterns do not match word-medial patterns. An example is *nostalgic* in British English, cited in Wells (1990) and shown in (58).

(58)	Syllabification	Problem
	a. [nɒ.stæl.ɔ̃ɪk]	Violating LOF: no word-final lax [ɒ]
	b. [nɒs.tæl.ɔ̃ɪk]	Violating LOI: no unaspirated word-initial [t]

In (58a), there is a syllable-final lax vowel, which is not found word finally. In (58b), there is a syllable-initial unaspirated [t] before a stressed vowel, which is not found word initially. Thus, neither analysis fits word-edge patterns. The word *gestation* poses the same dilemma ([ɔ̃ɪs.tei.ʃn̩] vs. [ɔ̃ɪs.tei.ʃn̩]). Such examples show that sounds at word-medial syllable boundaries may be subject to different conditions from those at word-edges. It also shows that, if *nostalgic* is syllabified as [nɒs.tæl.ɔ̃ɪk], the lack of aspiration in [t] is not triggered by an initial [s] in the same syllable. Therefore, in a word like *stay*, [s] need not be in the onset.

### 3.5.7. Summary

In (59) I compare four approaches to syllabification with regard to whether the rhyme exceeds VX (VX+), whether there is ambiguity in syllabification, and whether word-medial syllable boundaries agree with word-edge patterns.

(59)		Ambiguity	VX+	Edge patterns
	LOI and LOF	yes	?	yes
	Max stress	yes	yes	yes
	Max onset	no	no	no

WSP	no	no	yes
-----	----	----	-----

The Law of Initials and the Law of Finals are based on word-edge patterns, but there is ambiguity in syllabification (e.g. *whis.key* and *whisk.ey*) and rhymes can in principle exceed VX (e.g. *whisk.ey*), although additional rules can be proposed to limit them to VX (Vennemann 1988). The approach that maximizes stressed syllables also has ambiguity in syllabification (e.g. *Can.ad.a* and *Can.a.da*) and allows rhymes to exceed VX (e.g. *whisk.ey*). The Maximal Onset analysis is unambiguous and would usually not produce large rhymes, but it allows lax vowels to occur at the end of a syllable (e.g. *le.mon*), which does not match word-edge patterns. The approach using the Weight-Stress Principle is the most desirable in all respects.

### 3.6. The *a/an* alternation and the linking [r]

The English article is *a* before a C-initial word and *an* before a V-initial word. One might argue that the alternation supports the need for syllables to have an onset, as shown in (60).

#### (60) The [n]-insertion analysis of *a/an* alternation

[n] added as onset:     *an apple, an attack*

[n] not needed:        *a pear, a potato*

However, there are three problems with the analysis. First, if [n] is added to provide an onset for the following V, why does the article itself not need an onset? Second, why is [n] not added in other contexts, such as *be\_on time, go\_again, buy\_apples, panda\_act*, and *three\_Iranians*? One might suggest that there is a glide [j] after *be*, *buy*, and *three*, a glide [w] after *go*, and an epenthetic [r] after *panda*, and these can serve as the onset of the following vowel. However, why is [r] not used after *a*, such as *a [r] apple*? And why is [j] not used before [i], such as *a [j]*

*Iranian*? A third problem is that the original form of the article is not *a* but *an*, which comes from ‘one’ etymologically. Therefore, what we need to explain is not the insertion of [n] before V but the deletion of [n] before C, which has nothing to do with the need for an onset.

Given the etymological fact that the original form of the article is *an*, we can explain its alternation with the Weight-Stress Principle. Since *an* is unstressed, it should be light and cannot keep its coda [n]. The [n] can be moved to the next word, or else it can be deleted. The analysis is shown in (61).

(61) The [n]-deletion analysis of *a/an* alternation

	[n] moved	[n] unmoved	[n] deleted
<i>an apple</i>	[ə][næp][l]	*[ən][æp][l]	*[ə][æp][l]
		Unstressed VX	No need to
<i>a bee</i>	*[ə][nbi:]	*[ən][bi:]	[ə][bi:]
	Illegal [nb]	Unstressed VX	Best option

Both the insertion analysis and the deletion analysis assume that syllabification can occasionally cross a word boundary (especially when one word is completely unstressed). In addition, they both predict that *an aim* and *a name* should sound the same. The prediction is correct in casual speech but not in careful speech (Halle 1972, Kahn 1976, Kiparsky 1979). Regardless of how careful speech is analyzed, there is no obvious advantage in assuming the [n]-insertions analysis or the onset requirement.

A similar case of relevance is the linking [r] in English, such as *better* [r] *off* and *for* [r] *it*, and the intrusive [r], such as *law* [r] *and order* and *idea* [r] *of*. It may seem that [r] is added because syllables need an onset. On the other hand, the linking [r] is underlyingly present and so

its variation is a matter of deletion rather than insertion. In particular, for speakers who do not use a word-final [r], it can be moved to the next word, unless the next word cannot accept it. This is shown in (62).

(62) The [r]-deletion analysis

	[r] moved	[r] unmoved	[r] deleted
<i>for us</i>	[fo][rʌs]	*[for][ʌs] Final [r]	*[fo][ʌs] No need to
<i>for her</i>	*[fo][rhɜː] Illegal [rh]	*[for][hɜː] Final [r]	[fo][hɜː] Best option

The intrusive [r] can be analyzed as the result of analogy; for example, from [lɔ:] *lore* as to [lɔ:(r) ənd] *lore [r] and*, one gets [lɔ:] *law* as to [lɔ:(r) ənd] *law [r] and*. Therefore, what we can conclude is that syllables can have an onset, but they do not require one.

### 3.7. Epenthetic vowels between consonants

In some languages, epenthetic vowels can appear in consonant clusters. For example, while Dell and Elmedlaoui (1985) argue that Berber words can be made entirely of consonants, Coleman (2001) argues that they contain epenthetic vowels that serve as the syllable nucleus.

Similarly, the Tibeto-Burman language Jiarong reportedly has unusual onset clusters, yet Lin (1993) points out that such clusters may contain epenthetic vowels. For example, [kə-n-mʈʰə] ‘early’ can be pronounced as [kə-nə-mʈʰə]. In the former case, the syllabification seems to be

[kən][mʈʰə], which contains an unusual onset cluster [mʈʰ]. In the latter case, the syllabification could be [kə][nəm][ʈʰə], and there is not unusual onset cluster.

Polish is also known for having unusual initial clusters, such as [lɡɲ] in *lɡɲac* ‘cling to’ and [mkn] in *mknac* ‘scurry’. However, in the pronunciation of Sylvia Suttor, a native speaker of Polish, such clusters seem to contain epenthetic vowels, namely, [ləɡəɲ] and [mækʰɲ] respectively. As a result, *lɡɲac* sounds like three syllables to English speakers and *mknac* sounds like two syllables. Clearly, the presence of epenthetic vowels adds different possibilities for the analysis of syllable structure.

In the present study, I shall discuss one language that has epenthetic vowels, which is Jiarong in Chapter 11, and show how they affect the analysis of syllabification.

### 3.8. Are there parameters for the maximal syllable size?

Most phonologists believe that the maximal syllable size can vary from language to language. In generative phonology, the standard approach to variation is to propose a set of parameters (Chomsky 1981). If we can determine the parameters and their possible values, we can predict how many possible human languages there are and what properties they have.

Blevins (1995) proposes six binary parameters for the maximal syllable size, rephrased in (63), along with the values (or ‘settings’) for English. Similar parameters have been proposed before by Clements and Keyser (1983: 28-30).

#### (63) Binary parameters for the maximal syllable size (Blevins 1995: 219)

Parameters	English settings
Can the onset contain two sounds?	Yes

Can the nucleus contain two sounds?	Yes
Is the coda allowed?	Yes
Can the coda contain two sounds?	Yes
Can extra C occur initially?	No (Yes)
Can extra C occur finally?	Yes

The two parameters for the coda are dependent and yield three choices: no coda sound, a single coda sound, and two coda sounds. The other four parameters are independent. Therefore, there are 48 possible kinds of maximal syllable structures.

In Blevins's proposal, to have the simplest syllable, a language would make a 'no' choice for all the parameters, and to have the most complicated syllable a language would make a 'yes' choice for all parameters. In this regard, English has one of the most complicated syllable structures. In particular, if we consider the initial [s], as in *split*, to be an extra C in English, then English has made 'yes' choices for all parameters. Such a language would allow CCVVCC syllables in non-edge positions and CCCVVCCC syllables in monosyllables. In fact, Clements and Keyser (1982: 32) assume that C and V can each repeat at least three times, so that a maximal possible syllable is at least CCCVVVCCC.

The CVX theory predicts that such super-sized syllables are not only impossible but unnecessary. In particular, I shall show in Chapter 8 that word-medial syllables in English are limited to CVX, and apparent CC clusters are in fact complex sounds. In addition, extra Cs at word edges are not the result of parameter choices for the maximal syllable size, but the result of morphology, namely V-final prefixes support an initial C, V-initial suffixes support a final C, and consonant affixes constitute additional Cs. If other languages can be analyzed in the same

way, then there is no need to assume parameters for the maximal syllable size. Rather, all languages have the same maximal syllable size, which is CVX.

One might note that if all the parameters are set to ‘no’, the maximal syllable size is CV. Blevins (1995) cites the American Indian language Cayuvava as such a case, although it has some CC onsets, such as [tr, pr] (Key 1961), which in the present analysis can be seen as complex sounds. Let us suppose then that there are languages whose syllables are made of CV; in fact, I shall argue in Chapter 6 that Shanghai Chinese is such a case. Would we then need a parameter for the maximal syllable size, namely, CVX vs. CV? I suggest the answer be no. The reason is that when a CV syllable is stressed, it is likely to be lengthened, so that it is realized as [CV:] or CVV. This is true in Shanghai (Zhu 1995), and probably true in Cayuvava as well. Therefore, unless there is phonetic evidence to the contrary, there is no need to assume that the maximal syllable can be smaller than CVX. I shall discuss CV syllables again in Chapter 12.

### 3.9. The ‘spotty-data’ problem

I use the ‘spotty-data’ problem to refer to the fact that there are often not enough data for making reliable generalizations, even if we examine the entire lexicon of a language. To see the problem, let us consider the ratios between possible words and actual words in American English. The data on CVC and CVCVC words are shown in (64).

#### (64) The ‘spotty-data’ problem in English

Word form	Possible	Used	% used
CVC	2,415	615	25.5%
CVCCVC	5,832,225	6,000	0.1%

Excluding affixes and homophones, English has about 3,000 uninflected monosyllabic words (see Chapter 9), which include CVVC (842), CVC (615), CCVVC (453), CCVC (326), etc. The second most frequent type, CVC, includes 615 syllables. Now given 24 consonants (excluding [ŋ] from the onset and [h, tr, dr] from the coda) and 5 short (lax) vowels, there are  $23 \times 5 \times 21 = 2,415$  possible CVC syllables. This means that just one fourth of all possible CVC words are used. In dialects that have more short vowels, the percentage of occurring syllables could be even lower. If we consider disyllabic words, the percentage of occurring syllables becomes diminishingly small. For example, if any two CVC syllables can form a disyllabic word, there are about 6,000,000 possible disyllabic words, yet English only uses around 6,000 uninflected disyllabic words. This means that just 0.1% of all possible disyllabic words are used.

Why are so many possible words not used in English? One might suspect that there are phonological constraints that rule out most of the disyllabic words, but there are no known phonological constraints that would rule out 99% of the disyllabic combinations. Rather, the real answer in my view is that a language simply does not need many morphemes, which make up words. In particular, I shall discuss in Chapter 9 that English and Chinese use about 10,000 morphemes each, and many of them are infrequent. If a language only needs 10,000 morphemes, they often constitute just a very small fraction of possible words.

So if a language only needs 1% (or a few percent) of all possible words, which ones would be chosen? There are two possibilities: either the words are chosen more or less arbitrarily, or they are chosen according to phonological principles. It is often possible to look at the lexicon of a language and make various phonological generalizations. But without knowing whether a lexicon is an arbitrary or systematic collection of words, we cannot be sure whether the generalizations are real or merely artifacts.



In Chinese, whose morphemes are mostly monosyllables, the ratio between actual words and possible words is higher (see Chapter 5). Still, actual words only constitute a small fraction of possible words. Therefore, whatever patterns one might see in Chinese, questions can be raised. For example, Yip (1988: 82) suggests that Cantonese Chinese has a restriction against syllables that have two labial consonants, one in the onset and one in the coda, such as [pim] and [map], but she notes a few exceptions, such as [pəm] ‘pump’. Now, should we say that Cantonese disallows syllables with two labial consonants, or should we say that Cantonese happens not to have used any (or many) such words, but is in principle open to their use? The judgment of native speakers on non-words does not help much either, because it might simply reflect whether a word is or is not in the language, or how similar a non-word is to an existing word, or how many existing words a non-word is similar to, rather than what phonological principles are.

The spotty-data problem shows that it is often difficult to obtain true phonological generalizations and that sometimes we may not be sure what kind of rules a language has.

### 3.10. Summary

When a language seems to have large syllables, it is often because there are extra consonants at word edges. If we assume that all consonants must be syllabified, we have to assume very large syllables, such as CCCCVC (Hooper 1976a: 229) or CCCVVVC (Clements and Keyser 1983: 32), and then face the problem of over-prediction, because we cannot explain why non-edge syllables are much smaller. On the other hand, if we exclude extra consonants at word edges, we can maintain a smaller and consistent syllable size for both edge and non-edge

positions, but we have to explain why extra consonants can occur at word edges, a question that has not been satisfactorily answered before.

The CVX theory makes two claims. First, the maximal size of non-edge syllables is CVX (either CVC or CVV). Second, extra consonants at word edges are predictable from morphology. In languages with V-initial suffixes, an extra C is allowed in word-final position, because it can serve as the onset of the suffix V. Similarly, in languages with V-final prefixes, an extra C is allowed in word-initial position, because it can serve as the coda of the prefix V. Moreover, if a language has consonant affixes, they can be added regardless of whether they fit into a syllable. The claims make empirical predictions, which will be tested in the following chapters.

The concept of complex sounds (Chapter 2) explains how many underlying sounds can fit into each of the three CVX slots. In the extreme case, six underlying sounds can fit into a CVX syllable. An example is shown in (65).

(65)	Word	Sounds	CVX
	<i>prints</i>	[prints]	[p <sup>r</sup> ĭ t <sup>s</sup> ]

The word *prints* has six underlying sounds, which can merge into three complex sounds, where [p<sup>r</sup>] is formed from [p] and [r], the nasalized vowel [ĭ] is formed from [ɪ] and [n], and the affricate [t<sup>s</sup>] is formed from [t] and [s].

A critic may point out that, if more than three sounds can fit into a syllable, why do we not say that a syllable can contain more sounds, such as six? The answer is that not all sound combinations can fit into a syllable. The CVX theory predicts that only those combinations that can reduce to three or fewer complex sounds can fit into a syllable, and we shall see evidence in

the following chapters. In contrast, other theories need to explain why most combinations cannot fit into a syllable except those that can form three or fewer complex sounds.

I have also discussed several approaches to syllabification. I have argued that there is little evidence for a syllable to require an onset, either word initially or word medially. Instead, two independent principles—the Weight-Stress Principle and the VX limit on rhyme size—suffice to yield unambiguous syllabification and a satisfactory account of allophonic variations and word-edge effects.

To test the CVX theory, we should ideally look at every language and examine whether every non-edge syllable is within CVX and whether every extra C at word edges is predictable from morphology. As a beginning step, we can start with a selection of languages for which there is a sufficient amount of data on both the phonology and the morphology. In the following chapters I shall analyze five languages in detail: Standard Chinese, Shanghai Chinese, English, German, and Jiarong. The first two are chosen because they have a small syllable inventory, which we can examine exhaustively in order to determine the constraints that govern syllable-internal sound combinations. In addition, a minimal difference in the syllable coda between Standard Chinese and Shanghai Chinese has led to a dramatic difference in tonal behavior, which I shall discuss. English and German are chosen for their large consonant clusters, which in previous analyses are thought to require extra large syllables. Jiarong is chosen for its large initial consonant clusters, which are hard to account for in previous approaches. I shall argue that the analyses of these languages provide concrete evidence for the CVX theory.