J. Linguistics **48** (2012), 685–726. © Cambridge University Press 2012
doi:10.1017/S0022226712000205 First published online 26 July 2012

REVIEW ARTICLE

Invariant syllable skeleton, complex segments and word edges[1]

TOBIAS SCHEER

Université de Nice – Sophia Antipolis, CNRS 7320

San Duanmu, *Syllable structure: The limits of variation*. Oxford: Oxford University Press, 2008. Pp. xv + 275.

San Duanmu's *Syllable Structure: The Limits of Variation* raises a number of questions that are of general interest for phonological theory. Of special interest here are: the genesis and management of linearity in complex segments, the place of analogy (or paradigm uniformity) in grammar, the role of morphology in accounting for phonological patterns, the balance of static (distributional patterns) and dynamic (phonological processes) evidence for syllable structure, the role of stress in syllabification, and the import of corpus-based data for phonological analysis. In each case, Duanmu's proposals are evaluated according to their intrinsic consistency, the empirical record and the relevant body of literature. Alternative ways of handling the phenomena are offered, and these are fairly traditional in most cases. Duanmu's book is particularly relevant in the current constitution of the field where the see-saw movement between computation and representations seems to swing back in direction of the latter after having long been immobilised on the computational end. Standing clearly on the representational side, the theory exposed in the book aims to show that all surface strings may be reduced to a fixed and invariant syllable template, C(onsonant)V(owel)X. This enterprise is interesting especially in presence of another representationally-oriented theory, CVCV (Lowenstamm 1996, Scheer 2004), which also aims at reducing surface variation to an invariant syllabic skeleton, made of a monotonic sequence of CV units. However, the CVX and the CVCV templates are quite distinct, and the strategies that are used in order to accommodate the surface string are opposite (shrinking in the former case, expanding in the latter).

1. Introduction

San Duanmu's *Syllable Structure: The Limits of Variation* raises a number of questions that are of general interest for phonological theory. The pages

[1] This article was greatly improved by comments from Andrew Nevins and three anonymous *Journal of Linguistics* referees, all of whom I am indebted to. Many thanks also to Ewa Jaworska for thorough, competent and efficient proofreading.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

below address selected items, namely the genesis and management of linearity in complex segments (Sections 3 and 4), the place of analogy (or paradigm uniformity) in grammar (Section 5), the role of morphology in accounting for phonological patterns (Section 5.2), the balance of static (distributional patterns) and dynamic (phonological processes) evidence for syllable structure (Section 6), the role of stress in syllabification (Section 4.5), and the import of corpus-based data for phonological analysis (Section 7). In each case, Duanmu's proposals are evaluated according to their intrinsic consistency, the empirical record and the relevant body of literature. Also, alternative ways of handling the phenomena are pointed out, and these are fairly traditional in most cases.

The overall result is that while Duanmu's C(onsonant)V(owel)X theory raises important issues, it does not appear to be workable for it appears to be flawed in its internal logic in a number of respects, it does not cover the empirical record available or the relevant literature, and it does not explore the predictions made.

Nevertheless, I believe that Duanmu's book is relevant in the current constitution of the field where the see-saw movement between computation and representations seems to swing back in direction of the latter after having long been immobilised at the computational end.

Anderson (1985) uses the (im)balance between computation and representations as a prism in order to look at the history of phonology in the 20th century. He detects a regular see-saw movement between theories that stand far on one side of the spectrum, and others that approach the opposite extreme. Also, both aggregate states of phonology observe a reverse proportional relationship: when one goes up, the other goes down. Writing in the first half of the 1980s, when autosegmental representations were blooming and expected to be able to take over a good deal of the SPE-grounded rule overheat, Anderson (1985: 350) extrapolates that phonology stands at the dawn of a new computational, hence anti-representational, round. This was a correct prediction as the advent and evolution of Optimality Theory (OT) in the 1990s would show. Against the backdrop of the *Cambridge Handbook of Phonology* that Paul de Lacy edited in 2007 (de Lacy 2007), I have discussed the computational–representational balance in Scheer (2010a; see also Scheer 2011), arguing with Anderson (1985) that a sound phonological theory needs a theory of both aspects.

Duanmu's book clearly stands on the representational side: it exposes a representational theory of syllable structure, the CVX approach. The claim is that syllable structure in all languages is invariable (this is where the subtitle of the book comes from): there is only one syllable type, CVX (whereby C and X may be absent). Duanmu sets out to show that all surface variation, namely all consonantal and rhymal strings that look bigger than just C and VX, may be reduced to the CVX template. Blevins (2010: 287) correctly points out that this approach pursues the same goal of invariant

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

cross-linguistic syllable structure as CVCV (or strict CV), which is the framework that I am working in (Lowenstamm 1996, Scheer 2004, Szigetvári & Scheer 2005, Cyran 2010, among others). While the goal is shared, the solutions advocated by both theories are the exact opposite of one another: while CVX proposes to package surface clusters so that they can hold in single timing units (C, V and X), CVCV on the contrary expands syllabic constituency in order to accommodate the surface string. That is, CVX adapts the string of segments to a fixed syllable template while CVCV does the reverse: on the basis of minimal syllabic building blocks, CV units (onset–nucleus pairs), it adapts syllable structure to the linear string by allowing for empty constituents. For instance, the English word *trick* is parsed as $[[tr]_C$ $[[i]_V[k]_X]]$ in CVX, but as [tørikø] in CVCV (where 'ø' = empty nucleus). As can be seen, [tørikø] does not have any internal constituent structure and is therefore unable to express syllabic distinctions (such as closed vs. open syllable, onset vs. coda). This is precisely the idea of Government Phonology in general (Kaye, Lowenstamm & Vergnaud 1990) and CVCV in particular: syllable structure is defined by a system of lateral relations (of dependency type: government and licensing), rather than by constituent geometry.

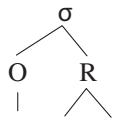
In sum, then, the present review article is motivated by the interesting constellation of two opposite means to achieve the same goal (invariance of syllable structure), and by the debate regarding the computational and representational balance in phonological theory.

2. How CVX works: A restrictive syllabic template and two tools for accommodating extra items

The formula CVX is a good shorthand: it displays two core properties of Duanmu's theory. On the one hand, there is the weight-based insight of moraic theory according to which the vocalic centre of the syllable can accommodate one additional item to its right, which may be either vocalic or consonantal: X can be a C or a V. On the other hand, the formula CVX indicates that this is as big as a syllable can grow: just one C, one V and one X – additional material must be accommodated by some other means. As noted above, this restriction is universal and motivates the subtitle of the book 'the limits of variation': the variation referred to here is the one regarding syllable size and syllable types. Describing the goal of the book, the author writes that '[f]or the most part, I shall focus on one theoretical question: What is the maximal syllable size in human languages?' (5), and this is indeed what the book is all about: page after page, language after language, everything is geared towards showing that however large (consonantal) strings can get (especially at word edges), they always fit into CVX. Thus, the cross-linguistic variation on syllable types and syllable size is zero – all syllables in all languages of the world are CVX (where C and X are optional).

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

The critical question, then, is how C, V and X are defined: what exactly can be accommodated by these symbols? This is the heart of CVX, and the subject matter of Chapter 2 of the book. The three items C, V and X are x-slots in the regular autosegmental sense, i.e. timing units (Clements & Keyser 1983, Levin 1985). Universally, a syllable can thus maximally accommodate three timing units, and CVX is shorthand for the structure in (1) (adapted from page 9).



onset/rhyme

timing slots/mora

C V X shorthand

At face value, this appears to be much too restrictive: what about branching onsets or super-heavy rhymes? Duanmu's struggle all through the book will be to explain that this restrictive structure is indeed able to accommodate all the apparent extra complexity that phonologists are familiar with.

In order to show that this is workable, the author devises two central tools: (i) his view on what is classically known as contour, or complex segments, and (ii) his view on analogy, or paradigm uniformity, which is called anti-allomorphy in the book. The way the author puts this (e.g. on page 4) is that extra consonants which occur at word edges are 'accounted for by morphology'. Let us first consider the former issue. The discussion of Duanmu's ideas in Section 4 below is preceded by a review of the relevant literature in the following section.

3. The analysis of sequences of acoustic events that count for a single segment

3.1 *Linear order: Lexical, phonological and/or phonetic representation*

Sequences of linearly ordered acoustic events which, however, appear to instantiate only one single phonological unit have left a noticeable footprint in the literature. Early discussion includes Martinet (1939) and Hoard (1967: 34ff.). The latter offers a solution that complies with the contemporary view on melodic organisation: like all other segments, complex segments are a feature matrix, which, however, is multi-columned. In the case of an affricate, the features [−cont] and [+cont] appear in different columns within the matrix (see also Hoard 1978). The next step were autosegmental representations, which were developed since the mid-1970s. They allowed for a straightforward formulation of this insight: Clements & Keyser (1983: 34f., 85ff.) set the standard autosegmental representation of affricates

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

as distinct stop and fricative branches that are dominated by a single skeletal position.

A different take first appears in Fujimura & Lovins (1978), who advance the idea that acoustic events which occur in sequence may be unordered in the phonology and linearised only upon phonetic interpretation (in their view, by phonetic realisation rules). While this concerns what they call the syllable core (a vowel plus a consonantal satellite on each side), the idea is adapted to (light) diphthongs by Kaye (1985: 289 – '[b]y universal convention the less sonorous of the two elements associated to the same point is produced first in the speech chain'), and to consonantal complex segments by Sagey (1986: 52ff.). The latter author distinguishes between two categories: contour segments (like affricates and prenasalised stops) whose linear order is specified in the phonology, and complex segments (like labio-velars: kp) which are unorderd in the phonology and receive their linear structure upon phonetic interpretation. Sagey's diagnostic for distinguishing the two types of monosegmental items are what she calls edge effects, i.e. the behaviour of the linearly distinct components in regard of phonological processes: affricates behave as stops in rules sensitive to their left edge, but as fricatives in rules sensitive to their right edge. Hence the order of the two branches must be available in the phonology. By contrast, labio-velars, for example, show no sensitivity to either edge: they always behave like both labials and velars. Therefore Sagey concludes that their linear order is absent from the phonology. This echoes a point made by Anderson (1976: 343): 'the segmental idealization is no longer quite so pervasive, since segments have internal structure which may be manipulated by rules'. In Anderson's data, phonological processes could concern either the nasal or the non-nasal part of a prenasalised consonant.

Pace Sagey, Lombardi (1990) argues that there are no edge effects: Sagey's data are either reanalysed or interpreted as the result of phonetic, rather than phonological, activity. Lombardi proposes that affricates (and only affricates: she is not concerned with other types of segments) are monosegmental items which bear a stopness and a continuancy feature that are unordered in the phonology. Linear ordering occurs upon phonetic interpretation for universal reasons (there are no mirror-image affricates like st, ʃt, etc., and there is no contrast between affricates and their hypothetical mirror images), and edge sensitivity is thus predicted to occur in phonetic processes. In addition, Lombardi levels anti-edge effects against phonological ordering, for example, cases where a rule applies AFTER stops and affricates. This shows that the phonology has access to both the continuant and the non-continuant component of affricates.

Also engaged in the debate on whether or not melodic primes are ordered in complex monosegments is Rennison (1998), who views the component that is realised last in two-membered complex items as floating, or 'lazy'. That is, the component that is first pronounced is the 'regular' consonant, which is

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

then modified by a new melodic prime, the 'lazy' one, that may either be lexically present or acquired from context (e.g. in assibilation). In Rennison's perspective, the linear order is thus programmed in the phonology: the component that is realised first is specified as such in the lexicon, and may be supplemented in further pronunciation either by other melodic primes that are lexically present (floating), or that are provided during phonological computation.

The study of complex segments in general, and the view that features are unordered in particular, has also consequences for the kind of melodic primes that are used: monovalent, rather than binary features are necessary in order to run unordered complex segments. An early contribution in this direction is based on the study of prenasalised consonants: Anderson (1976: 332) argues 'against any solution to the problem of prenasality in which a single feature has the entire segment as its domain' because the difference between N, ^{n}C and C^{n} is one of timing, i.e. the moment at which the velum is lowered. 'For this reason any feature which purports to characterise the entire segment uniformly, like [continuant] or [sonorant], is apparently destined to fail'. Lombardi (1990: 377f.) discusses the same issue from a different perspective: how could [−cont] and [+cont] occur simultaneously in the same segment, i.e. an affricate in her case? On the basis of other equipollent features such as [±anterior], Lombardi concludes that opposite values of the same feature cannot co-occur within a single segment. She therefore introduces two distinct features that do the same labour as [−cont] and [+cont]: [stop] and [cont]. An affricate bears both of them (in unordered fashion), and as a consequence both must be monovalent, since [stop] is the equivalent of [−cont].

Steriade (1993: 461; 1994: 210ff.) arrives at the same conclusion for different reasons, namely the fact that the only reliably attested segmental contours – for nasality and continuancy – occur with released plosives, and also that plain consonants, i.e. unaspirated and unglottalised, are always good landing sites for aspiration and glottalisation. If they bore [−spread] and [−constricted] specifications, they would be expected, at least in some cases, to block the spreading of the positive value of these features (since positive and negative values of the same feature cannot co-occur). Against this background, Steriade (1993: 459f.; 1994: 216f.) introduces both universal and parametric conditions for the co-occurrence of features within a single segment. Universal in her view are physiological impossibilities such as the simultaneous realisation of an aspirated and a glottalised release: [+spread] and [+constricted] cannot be simultaneously pronounced. On the other hand, languages choose parametrically which features are allowed to hook onto a single segment. Steriade (1994: 217) therefore argues that 'even features that are not simultaneous on the surface must pass a test of mutual compatibility in order for the structure encompassing them to count as monosegmental'.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Underlying this perspective is the idea that some feature combinations within a single segment are uninterpretable, either universally or on a language-specific basis. There are two ways to implement this idea, though. In the one we have seen, uninterpretable feature combinations are illegal and cannot exist in the system. An alternative view is that the uninterpretability of a given feature set causes a linear differentiation upon phonetic interpretation: the features are uninterpretable SIMULTANEOUSLY, but can be pronounced in a sequence. Based on Jakobson, Fant & Halle (1952), this line of thought is followed, for affricates, by Hirst (1985: 93), Hualde (1988), La Charité (1993), Rubach (1994), Kim (1997), Clements (1999) and Kehrein (2002: 7ff.). For these authors, there is no [+cont] component in affricates at any derivational stage, or in the lexicon; affricates are regular stops whose only peculiarity is to possess a [+strident] feature. Since [−cont, +strident] cannot be executed simultaneously, the segment is spelled out as a sequence of acoustic events at the phonetic level. With reference to feature co-occurrence in (English) onset clusters, Hirst (1985: 92) extends this perspective to (English) muta cum liquida (pl, tr, gl, etc.), which according to him are a single segment that is specified for [−cont, +son, −nas]. Since no segment can be an oral sonorant stop, the features at hand cannot surface unless their components are separated in a linear order.[2]

There are thus two distinct views on the question why phonologically unordered features appear as a sequence on the surface. On one view, the reasons are purely phonetic (and universal) and have nothing to do with phonological specification; the stop component is pronounced before the fricative component (e.g. Lombardi 1990). On the other view, the reasons are also phonetic but anchored in the phonology – uninterpretability of feature sets stems from the fact that affricates are strident stops.

Let us now turn to the question of how complex segments arise (whatever their internal structure). Are they present in the lexicon, or are they distinct segments underlyingly which merge in the course of the derivation? That lexically distinct items merge into complex segments at least in some languages and at least for some types of complex segments is shown by the trivial fact that in many Bantu languages the nasal and the non-nasal component of prenasalised stops belong to distinct morphemes (e.g. Herbert 1977). The issue is extensively discussed by Steriade (1994), who concludes that some mergers are obligatory and universal, while others are governed by parametric variation. Aspirated stops C^h instantiate the former case: they may occur as such in the lexicon, but if a C and an h happen to form a

[2] A third view on linearisation is found in the above-mentioned article by Kaye (1985: 289), who argues that linearisation is sonority driven: 'By universal convention the less sonorous of the two elements associated to the same point is produced first in the speech chain'. While this works for affricates and light diphthongs, a special provision is needed for prenasalised consonants, for example.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

tautosyllabic cluster, they always merge. This is, Steriade (1994: 215) argues, the reason why no language opposes C^h and C–h. On the other hand, parametric variation is responsible for the fact that $n + t^h$ merge into a prenasalised stop $^nt^h$ in Tlacoyalco Popoloca (Oto-Manguean), but not in the closely related Western Popoloca (Steriade 1994: 272). Van de Weijer (1996: 158) argues along the same lines regarding prenasalised stops. Like Steriade, Hirst (1985: 92) holds that muta cum liquida are single segments in some languages (English in his case), and like her, he admits that the same sequences may be bisegmental in other languages (e.g. when they are broken up by vowels, as in Moroccan Arabic). At variance with Steriade, however, Hirst does not provide for a mechanism that merges lexically distinct segments into a single segment during phonological computation: his monosegmental muta cum liquida are already monosegmental at the underlying level.

3.2 *Diagnostics for identifying monosegmental items*

In this section we take a closer look at the kind of arguments advanced to distinguish monosegmental from bisegmental items. A very common argument concerns the distribution of complex segments, which is one of singletons, rather than of clusters: for example, complex segments may occur in positions where clusters are impossible (Anderson 1976: 331; Steriade 1993: 458; Clements 1999: 271; Hualde 1991).

Another common argument looks at inventories: for example, in a language where voiced prenasalised stops are the only voiced stops, their interpretation as a cluster would leave the language without voiced stops (that are independent of a preceding nasal) (Anderson 1976: 331). Along the same inventory-based line is what Clements (1999: 271) calls 'non-compositionality: a sequence of sounds, say a stop + fricative string, cannot be plausibly decomposed into a sequence of two independently-occurring phonemes'.

Contrast is also often quoted. For example, when prenasalised stops contrast with singletons (stops and nasals for instance), but not with clusters, one may conclude that they are monosegmental as well (Anderson 1976: 311). Clements (1999: 272) calls his argument based on contrast 'non-commutivity: the stop and/or fricative component [of an affricate] does not commute with another phoneme'. Still based on contrast, but this time with clusters, is what Clements (1999: 272) calls 'contrastivity: some languages (e.g. Czech, Polish) have lexical contrasts between affricates and the corresponding stop + fricative sequences'. The contrast found in Polish between *czy* [t͡ʃɨ] 'whether' and *trzy* [tʃɨ] 'three' (and a number of other minimal pairs of the kind) is often quoted when this point is made (e.g. Hoard 1967: 36f.; Anderson 1976: 311; Herbert 1977: 440; Ewen 1982: 52; Clements & Keyser 1983: 34f.).

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Another argument concerns syllabification and sonority sequencing. Word-initial prenasalised stops, for example, would violate the latter, were they separate segments (Ewen 1982: 42, 54; Hualde 1991; van de Weijer 1992: 133; 1996: 54). The corresponding argument that Clements (1999: 271) calls tautosyllabicity is about syllabification: 'affricates syllabify as syllable onsets in contexts where bisegmental stop + fricative sequences usually syllabify as coda + onset sequences'.

Another factor for deciding whether a sound sequence is mono- or bisegmental revolves around the aforementioned feature co-occurrence. Steriade (1994: 217) provides a definition of what a single segment is which relies partly on theory-internal considerations, but involves one theory-neutral factor: a sound sequence that has more than one place of articulation cannot be a complex segment. Hence muta cum liquida pr, kl, tr, etc. do not qualify. Note, however, that Steriade's notion of single segmenthood is more complex than that (and may appear inconclusive): she operates with two distinct notions, 'featurally monosegmental' and 'structurally monosegmental'. An item may be featurally non-monosegmental, but qualify for structural monosegmenthood: this is the case of muta cum liquida. According to Steriade (1994: 213, 221), these have indeed the same structure as simplex (released) stops (i.e. an A_0 followed by an A_{max} position in her system, where A stands for aperture position): they are structurally monosegmental. At the same time, though, muta cum liquida are bisegmental regarding their featural makeup, which contains two distinct sets of place features.

Among all arguments for monosegmenthood, though, the one drawing on phonological processes is certainly most pervasive. Sagey (1986: 69ff.) offers a number of relevant examples. One (73ff.) is built on Kinyarwanda (Bantu), where compensatory lengthening of the vowel preceding prenasalised (/ku-ngana/→kuu^{ŋ}gana 'to be equal'), but following labio-velar consonants (/ku-gu-ir-a/→kug^{w}iira 'to fall on') is observed. This is explained by the fact that two lexically independent items have merged into a complex segment, and the x-slot that was vacated by this operation is colonised by the adjacent vowel. Clements (1999: 272) contributes two patterns to this argument: 'inalterability: affricates fail to undergo rules of cluster reduction that apply to bisegmental consonant sequences; [and] inseparability: the stop + fricative sequence cannot be broken up by epenthesis, reduplication, speech errors, and so forth'. Other cases are quoted by Hualde (1991: 127) and van de Weijer (1996: 134ff.). The latter author mentions the Sahidien variety of Coptic, where the definite article appears as *p-* (masc.), and *t-* (fem.) when the noun begins with a single consonant (*p-to* 'the earth', *t-soone* 'the sister'), but is realised as *pe-*, *te-* when followed by a cluster (*pe-thab* 'the leaven', *te-hbuur* 'the left'). Like single consonants, the affricate t͡ʃ does not provoke the vocalisation of the article (*p-čo* 'the armpit', *t-čamee* 'the calm'), which shows that it counts as a single segment. Another source of

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

evidence is reduplication. Van de Weijer (1996: 135) reports a case from Ewe (Niger-Congo) where only the first consonant of a word-initial cluster is reduplicated (ɲ*ra* – ɲ*a*-ɲ*ra-la* 'to rave, raver', *fle* – *fe-fle* 'to buy, bought'). When the initial cluster begins with an affricate, the entire affricate is copied (*d͡zrá* – *d͡zá-d͡zrá* 'sell, selling'), and this is also the case when the affricate is alone (*t͡sí* – *t͡sí-t͡sí* 'to grow, grown'). Affricates are thus treated as single segments.

Steriade (1994: 204f.) argues against the bisegmental status of what is traditionally interpreted as a branching onset: 'one expects all prosodic constituents … to be the domain of phonological rule application. But the subconstituents of the onset … never play this role of prosodic domain'. McCarthy & Prince (1996: 70) make an argument from templatic morphology: '[i]n Modern Hebrew and in Arabic, the respective affricates *c* and *j* spread intact to multiple positions in the skeleton: *kicec* "he cut" … , *hajjaj* "he made the pilgrimage to Mecca". To derive these results it is crucial that affricates constitute a single melodic element that spreads' (Lombardi 1990: 409 and van de Weijer 1992: 133 also make this point).

Lowenstamm (2003: 340ff.) presents an argument along the same lines, which, however, offers a diagnostic for the mono- or bisegmenthood of the same phonetic object. Quadriliteral forms of Chaha (Ethio-Semitic) identify, among other things, by bearing 1 in V_1. Hence the regular form *mɪsäkär-ä* 'to testify, 3sg masc.' and *fɪcʹänäq-ä* 'to squash, 3sg masc.' (*ä* denotes a low schwa). In *qʷɪräcʹä* 'to take a handful and close the hand, 3sg masc.', however, one consonant seems to be missing. Lowenstamm argues that the palatal ejective stop *cʹ* may represent two distinct lexical items (and two distinct items in the consonantal inventory of the language): in *fɪcʹänäq-ä*, it is a monosegmental unit (/tʲʼ/), while it identifies as a bisegmental item (/tʼy/) in *qʷɪräcʹä*. Steriade (2007: 140) reports a number of parallel cases where a language contrasts a surface sequence XY according to whether it is mono- or bisegmental: kw vs. kʷ in reconstructed Indo-European, tˀ vs. the ejective tˀ in Yokuts (native American, California).

Lowenstamm (2003: 343ff.) further argues that muta cum liquida are never branching onsets in any language: branching onsets do not exist, and muta cum liquida are either monosegmental or heterosyllabic (like in typical Semitic languages). An argument he makes against branching onsets as such is their inertness with respect to stress: stress assigning algorithms may or may not be sensitive to branching rhymes, but never make a difference between simplex and what is supposed to be branching onsets.

On the phonetic side, Byrd (1996a, b) has found that the segmental unit displays relative stability of timing and coordination: there is less gestural variability and overlap (in timing) within segments than among sequences of segments. Also, Byrd (1996a: 161) has observed that speech rate has less impact on segments than on sequences of segments. Related to this is the obvious point that monosegmental items should be shorter than bisegmental

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

clusters (especially according to autosegmental standards where the former occupy one x-slot, while the latter extend to two timing units). The duration of familiar complex segments such as affricates, prenasalised stops, etc., however, is not significantly longer than the one of simplex consonants (e.g. Ewen 1982: 42, 60; Sagey 1986: 78; Maddieson 1989; van de Weijer 1996: 158f.).

Finally, Fromkin (1971: 33) makes a point regarding affricates on the basis of a corpus of English speech errors: '[i]t is interesting to note that while [str], [pl], [kr], [bl], [fr], etc., as well as final clusters, reveal the splitting of clusters into segments, not a single example in my own data, or the English examples cited by others, shows a splitting of [tš] or [dž] into sequences of stop plus fricative'.

3.3 *Diagnostics for identifying the internal structure of monosegmental items*

A related issue concerns the internal structure of items that are agreed to be monosegmental. It was reported in Section 3.1 above that there are two basic options: linear structure is either a phonological or a merely phonetic property. In the latter case, either purely phonetic reasons cause linearisation (e.g. aerodynamics), or linear realisation is a consequence of the attempt to execute features that cannot be simultaneously pronounced (for either universal or language-specific reasons).

Again, the evidence that is most pervasively used in order to distinguish between these options comes from phonological processes. Anderson (1976: 337) writes that '[i]f segments with such internal structure can exist, however, we would also expect phonological rules to be able to create them, and in general we would expect phonological rules to be able to manipulate the size of the domain of a feature specification wherever that is a variable'. He then examines a number of relevant processes, and concludes, on page 343, that 'segments have internal structure which may be manipulated by rules'. Recall from Section 3.1 that so-called edge effects, i.e. processes which involve only the left or right element of complex items, are the basis for Sagey's (1986) distinction between contour and complex segments. This distinction was then counteracted by Lombardi's (1990) anti-edge effects. Other cases where phonological processes are used as a diagnostic for the internal structure of monosegmental items include Rosenthall (1992: 252ff.), La Charité (1993), van de Weijer (1993: 89, 99), Rubach (1994) and Kim (2001).

Another recurrent argument is overgeneration. In order to avoid the possible existence of objects (and hence contrasts) that are unattested, Sagey (1986: 50) restricts branching within the feature geometrical tree to terminal nodes (which represents an evolution with respect to Clements' (1985) original feature geometric model). Clements (1999: 272) writes that '[a]ny further representational distinction, such as one involving a feature [+continuant] (whether contoured or simultaneous), would make incorrect predictions

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

concerning affricate behaviour, and massively overgenerate the number of potential contrasts predicted by the representational system', and concludes, on page 277, that 'special analyses of the sort represented in (1) [i.e. involving the feature [+cont]] are not just superfluous, but predict the existence of many more speech sounds than are actually attested'. Van de Weijer (1992: 133) and Rosenthall (1992: 250f.) are also concerned with overgeneration.

Another line of attack concerns the (universal or language-specific) interpretability of feature combinations that was already discussed in Section 3.1. On the one hand, the internal structure of monosegmental items is strongly marshalled by the universal impossibility to accommodate two conflicting values of the same feature in a single melodic makeup. Hence Lombardi (1990: 377f.), who argues that the correct representation of affricates involves the unordered cohabitation of [+cont] and [−cont], must express this distinction by two different features, [stop] and [continuant]. These are then necessarily monovalent. Steriade (1993, 1994) is also strongly committed to monovalent features, for the same reason. On the other hand, those who argue that linearisation into distinct acoustic events is a consequence of the obligation to pronounce simultaneously uninterpretable feature combinations in turn need to propose a specific featural content for monosegments that creates such a conflict. Hence the presence of [+strident] in affricates in Clements' (1999) view, which causes linearisation since it cannot be pronounced together with [−cont]. Finally, considerations regarding the general feature-geometric setup may also determine the internal structure of monosegments. For example, Rosenthall (1992: 250) rejects Sagey's above-mentioned ban on branching class nodes because prenasalised consonants need both a [+nas] and a [−nas] specification (he does not mention the monovalent way out that is commonly chosen, though).

Steriade (1994) makes a point based on the clustering possibilities of stops vs. non-stops. She first asks 'why is it that the only reliably attested segmental contours – for nasality and continuancy – are found among released plosives?' (page 210) and then argues that this is because released plosives have internal structure, i.e. a closure and a release POSITION. Against this background, she explains that the curtailed clustering possibilities of non-stops is due to their lack of the release position: they are mono-positional and therefore cannot accommodate extra structure related to nasality or continuancy: 'plosives have more clustering possibilities than continuants, because plosives are bipositional' (Steriade 1994: 203). The increased clustering possibilities of released stops thus motivate a bipositional internal structure of complex monosegmental items (see also Steriade 1993: 406).

Finally, van de Weijer (1993: 94; 1996: 138f.) argues that in affricates, the continuant part has a tighter relationship with the place of articulation of the whole than the stop component, and that this should be represented in

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

the feature geometric tree. His proposal is based on the 'observation that affricates are typically found at place of articulation where fricatives are also found' (van de Weijer 1993: 94).

4. Duanmu's complex sounds

The literature mentioned in the preceding section is by and large absent from Duanmu's book: the author develops the notion of complex sounds without referring to the arguments, claims and reasoning regarding this matter that were made in the past.

When Duanmu's ideas are exposed below, they will be systematically related to the relevant literature. The overall picture that emerges converges with Blevins' (2010: 287) appraisal in her own review of the book that appeared in *Linguistic Typology*: '[w]here surface syllables larger than CVX occur in word-medial position, complex onsets and trisegmental and longer rhymes are squeezed into single timing slots by a poorly justified process of complex sound formation'. Blevins also observes that 'even for the five languages for which a detailed treatment is provided, the evidence for (C)VX syllables over larger syllable types is weak, having more the feel of possible world treatments than empirical arguments' (Blevins 2010: 288). She adds, on page 289, that '[a] scientific approach to sound patterns demands that there should be some empirical test for single segment vs. cluster analysis, but none is provided in the volume'.

Note, however, that another review of the book, by Gregová (2009), is rather positive. Regarding the complex sound analysis of affricates, Gregová (2009: 93) writes, for example, that 'Duanmu's solution is lucid and clear: affricate sounds are made with the simultaneous features (or, as he calls them, GESTURES) [+stop] [+fricative]. This, I think, faithfully reflects the phonetic nature of affricate sounds (the combination of a closure and narrowing …)'.

4.1 *Complex sounds: Genesis and general properties*

In order to understand what a complex sound is, we must first understand what a 'sound' is in Duanmu's system: the author only talks about sounds, not about segments. On pages 14–17, two criteria are established that an item needs to meet if it wants to qualify as a sound: it must be independently pronounceable (hence a feature alone cannot be a sound), and it must be sequenced in time: simultaneously pronounced items cannot be part of different sounds.

Sounds are thus the minimal units that have linear status. On page 25, the author introduces complex sounds in the following manner: '[w]e can think of a complex sound as the merger of two (or more) sounds'. This indicates that complex sounds are absent from lexical recordings in

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Duamnu's system: the lexicon is a linearly ordered sequence of sounds. Unfortunately, though, and unlike other authors (e.g. Hirst 1985, Steriade 1994; see Section 3 above), Duanmu is not explicit on the issue, and the reader cannot be sure that this conclusion is correct. All sources that I have come across agree that sometimes complex segments are complex already in the lexicon, and there is also evidence for contrasting bisegmental CX vs. monosegmental C^X in the same language (e.g. Lowenstamm 2003; see Section 3.2 above). To the extent that Duanmu really implies that all complex segments are non-complex in the lexicon, we can note that he stands alone with this position.

In Duanmu's system, lexically adjacent sounds may thus combine into complex sounds (that take up the place of one x-slot in (1) above) by virtue of some computational action (which is not further specified) that occurs prior to the construction of syllabic constituents. This is parallel to Steriade's (1994) merger operation (but no reference to this article is made).

Like a number of other authors, Duanmu argues that muta cum liquida and other types of what is traditionally held to be branching onsets may be monosegmental.[3] As far as I can see, though, nobody has ever proposed that sequences which do not qualify for tautosyllabicity (according to regular sonority sequencing), i.e. strings of falling or equal sonority, may also be monosegmental (e.g. Steriade 1994: 215 strictly upholds the tautosyllabic condition for monosegmenthood). This is because whenever there is a diagnostic for closed syllables in a language, the strings at hand are heterosyllabic. Pace the literature and the evidence, though, Duanmu is also prepared to count sequences like *lp* as monosegmental. Issues raised by granting this option are discussed in the following section.

In sum, Duanmu's fusion operation is entirely unrestricted: it can in principle combine ANY NUMBER and ANY KIND OF ADJACENT SOUNDS into a complex sound (25–31) as long as Duanmu's (1994) No Contour Principle is satisfied (which is discussed in Section 4.3 below).

[3] A peculiar way to motivate affricate status of tr, dr in English is found on page 183: '[p]honetically [tʃ, ʤ, tr, dr] are all affricates, although [t͡ʃ ʤ] are often listed as single sounds but [tr, dr] are listed as consonant clusters'. Duanmu does not explain what a 'phonetic affricate' is (as opposed to a 'phonological affricate'?), nor which (presumably 'phonetic') properties allow us to identify them as such. Also, the reader wonders why tr, dr count as 'phonetic affricates', but other classical onset clusters with labials and velars (pr, br, pl, bl, kr, gr, kl, gl) are left unmentioned. Elsewhere the author is explicit on the fact that all items mentioned are complex sounds and hence affricates (174–175, see Section 5.2 below) – so is one to conclude that the sequences with labials and velars are complex sounds, but not 'phonetic affricates'? A *JL* referee suggests that maybe Duanmu opposes tr, dr to the other onset clusters because they are homorganic – but since the author is not explicit on this division, the reader is left guessing.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

4.2 *Linearity is purely phonetic, hence lexical linearity is neutralised (/lp/ and /pl/ = [pl])*

We have seen in Section 3.1 that Sagey (1986) distinguishes contour and complex segments according to linearisation: while the former have phonological ordering, the latter are only linearised during phonetic interpretation. Recall that according to Sagey (1986: 51ff.), contour segments include affricates, prenasalised consonants and contour tones, while sequences of stops with distinct place of articulation like kp, pt or tk^w instantiate complex segments (Sagey 1986: 57f.). Hualde (1988), Lombardi (1990) and Clements (1999) are among those authors who deny the existence of any phonological ordering of the components of affricates, hence making this type of item also a complex segment (in Sagey's classification).

As far as I can see, though, there is no attempt to interpret the other prominent contour segment, prenasalised consonants, as a phonologically unordered structure. There is good reason for the absence of unordered interpretations. Lombardi (1990: 377) explains that linearisation of affricates occurs post-phonologically for UNIVERSAL phonetic reasons: there are no mirror-affricates where the first component articulated is a fricative, followed by a stop. The same cannot be said for prenasalised consonants: their mirror image, postnasalised consonants, exist as well. Hence linearisation cannot be decided by some universal phonetic mechanism; rather, it needs to be specified in the phonological representation. Steriade (1994: 206) (who upholds phonological ordering also for affricates) is explicit on this. In her system where released stops are made of two internal positions, closure (C) and Release (R) (in this linear order), [nas] linked to both C and R produces a regular nasal (m, n), [nas] associated only to C results in a prenasalised stop (^{m}b), while [nas] linked only to R is realised as a postnasalised stop (b^m).

Without discussing or mentioning these issues and the literature at hand, Duanmu takes on the position that ALL complex sounds are unordered in the phonology.[4] Following Articulatory Phonology (Browman & Goldstein 1989), he thinks of complex sounds in terms of gestural overlap: 'all features in a complex sound are simultaneous, in the sense that no sequential timing difference can be used distinctively' (26). He also makes clear that linearisation is due to purely phonetic and universal properties: 'the articulatory gestures for an affricate are made simultaneously, but the phonetic results are sequential, owing to the nature of articulatory aerodynamics' (22). Hence Duanmu is not among those who hold that linearisation is done in the phonetics, but provoked by the cohabitation of a set of features that are uninterpretable simultaneously (e.g. Hirst 1985; Clements 1999, see Section 3.1).

[4] As will be discussed in the following section, a consequence of this move is the exclusion of prenasalised consonants from the set of possible complex sounds.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Given this context, here is the list of sequences that are analysed as complex sounds in the book: labialised velars k^w (25), labio-velars $\widehat{kp}$ (27), a vowel and a glottal stop V^{γ} (31), obstruent–sonorant clusters (branching onsets in traditional analysis) p^l, t^r, etc. (26, 50), but also the reverse items $^l p$, $^r t$ (50, which Duanmu says are rare, on page 26). Beyond things that appear as a linear sequence of two items on the surface, complex sounds may also involve three surface items, or only one. On Duanmu's analysis, English r is in fact r^w in syllable-initial position, and when involved in a branching onset a triple complex segment is produced as in *try*: $t^{rw}y$ (29). On the other hand, the simplex surface ʃ that may appear at the word boundary of *this year* is also a complex sound, made of s + yod (26), and nasal vowels $\tilde{V}$ may be an underlying VN sequence (30–31).

A consequence of Duanmu's claim that the components of complex monosegments are never ordered in the phonology is that lexical linearity is neutralised. Hence Duanmu makes the following point (about English): 'why is there no *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*' (50, emphasis in original).

This reasoning does not go through, though. In order to see that, let us first consider the situation in languages where word-initial #RT clusters do occur: many Slavic languages and occidental varieties of Modern Arabic are cases in point (Scheer 2007).[5] In Czech, for example, #kl- and #lk- not only both occur word-initially, but also form 'minimal pairs': compare *klát* 'to cut wood' with *lkát* 'to whine, to moan'. Following Duanmu's standards, one concludes that in such a language the R of word-initial #RT clusters cannot belong to the same syllable as the first vowel: it cannot form a complex sound with the obstruent since linearity in complex sounds is neutralised (/kl/ and /lk/ would both come out as [kl]). The only option for #RT clusters in Czech, then, is to be bisegmental: the #R remains unsyllabified. Due to anti-allomorphy (see Section 5.1), stray (or in traditional terms: extrasyllabic) consonants are legal as long as there are V-final prefixes (for extra #C) or V-initial suffixes (for extra C#). Unsurprisingly, Czech does have V-final prefixes, hence *lkát* can happily (or rather: must) identify as *l.kát* where the lateral remains unsyllabified. This is confirmed by Duanmu's analysis of word-initial clusters in Jiarong (Tibeto-Burman, Chapter 11), where large word-initial clusters occur that include #RT. Items such as #rpV, for example, are analysed as *r.pV* on the grounds of the aforementioned potential prefix-final vowel (226, 230).

The question, then, is why the initial #RT cluster of the putative English word /lpum/ could not be bisegmental along the same lines: English has

[5] Here and henceforth T is shorthand for any obstruent, R for any sonorant.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

V-final prefixes, hence an analysis in terms of an unsyllabified word-initial liquid, *l.pum*, is an option. Behind this issue is the question what it is that makes some languages tolerate #RT clusters, while others refuse to accommodate them. Recall that Duanmu's explicit answer for the absence of English #lp is that the only option would be a complex sound, which, however, would be pronounced #[pl]. This means that for some reason, which Duanmu does not discuss, the bisegmental (and heterosyllabic) option for #lp is barred in English. One guesses, then, that it is this prohibition for word-initial #RT clusters to acquire bisegmental status that drives the parametric variation at hand: #TR-only languages bar bisegmental status for #RTs, while languages where #RTs occur on the surface allow for this syllabification.

But in any case, the reason that Duanmu gives (50) for the absence of *lpum* in English cannot be the real reason, according to his own standards.

4.3 *The No Contour Principle and the exclusion of prenasalised consonants from the class of complex sounds*

It was mentioned in Section 3.3 that the study of complex segments has yielded consequences for their featural makeup, and for feature theory in general. Namely the impossibility for a segment without internal linear structure to accommodate the plus and the minus value of the same feature was the motor for introducing privative features: instead of being [+cont] and [−cont], an affricate in Lombardi's (1990: 378) analysis bears the monovalent features [stop] and [continuant]. The same insight was already made by Sagey (1986: 65): 'the absence of an articulation is never specified' (Sagey's solution were monovalent class (place) nodes, though). Steriade (1993: 406f.; 1994: 210ff.) is also along these lines.

In a sense, Duanmu's (1994) No Contour Principle (NCP), introduced on pages 25–31, is an implementation of this insight – but again the book offers no discussion of the relevant literature, or of eventual differences with previous formalisations.

Instead of prohibiting the presence of the same feature (with opposite values) within a segment (that lacks internal linear structure), Duanmu argues for a weaker version where several instances of the same feature are allowed within a (complex) sound – but cannot co-occur under the same articulator: 'an articulator cannot take the same feature (F) twice within one sound' (26). Articulators in Duanmu's system are similar to non-terminal nodes in a regular feature geometric tree: the six articulators assumed are Vocal-cords, Soft-palate, Tongue root, Dorsal, Coronal and Labial.[6]

[6] Unfortunately, there is no how-to that allows the reader to understand, for any given sound, which features exactly depend on which articulators. The internal structure of six consonants (p, m, s, ʃ, t͡ʃ l) and four vowels (a, i, ɪ, u) is made explicit on page 24, and a few

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Duanmu's example to show what the NCP rules out is the merger of k and m. 'Since [k] is [−nasal] and [m] is [+nasal], there is a conflict' (27): the complex sound must be either entirely nasal, or entirely non-nasal. Duanmu argues that k^m (k before m), ^mk (m before k) and km (k and m simultaneous) are ruled out by the NCP since they involve simultaneous [+nasal] and [−nasal] action of the articulator Soft-palate. Were [+nasal] and [−nasal] realised by two distinct articulators, though, the result would be a viable complex sound. For example, if [+nasal] depended on the Labial articulator, and [−nasal] were dominated by Soft-palate, k and m could engage into a legal complex sound. The fact that km is declared illegal by Duanmu thus allows the reader to conclude that the feature [nasal] cannot depend on any other articulator than Soft-palate. This is never discussed or made explicit in the book, but it makes physiological sense since the velum is located in the area of the soft palate.

(Complex) sounds that bear two instances of the same feature are thus legal when the instances at hand either occur under two distinct articulators, or have the same value. Examples are k^w ([+stop] is realised by Dorsal and [−stop] by Labial (25)), kp ([+stop] is realised by Dorsal and Labial (25)) and ʃ ([+fricative] is made by both Coronal and Dorsal (25)).

Thus, given the NCP, affricates could not be complex sounds when standard features are assumed, since they have only one place articulator: [+cont] and [−cont] would be realised by the same articulator, Coronal, for example, in the case of t͡s. It was mentioned at the outset of this section that this is exactly the issue that Sagey (1986), Lombardi (1990) and other authors were struggling with: their solution was to introduce two distinct features for the description of the stop and fricative components of affricates, [stop] and [continuant], and to make them privative. One senses that this is also what makes Duanmu adopt non-standard features: t͡s, for example, is made of the Coronal articulator, which dominates [+stop] and [+fricative]. But again there is no discussion in the book where these choices would be explained, and the reader is not referred to any literature either, despite the obvious precedent.

Finally, a direct consequence of Duanmu's feature-geometric setup, and also of the fact that linearity is banned from complex sounds (see Section 4.2 above), is that prenasalised consonants, an archetypical complex segment in the literature, are excluded from the set of possible complex sounds: they would involve [+nasal] and [−nasal] under Soft-palate. This is quite dramatic a departure from several decades of study of prenasalised consonants and the massive evidence that speaks in favour of their status as complex

more identities are provided on the following pages. But in absence of the key to the distribution of features over articulators, the predictions of the system (i.e. which complex sounds should be possible and which ones are ruled out) cannot be computed. Blevins (2010: 289) makes the same observation.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

segments (see Section 3.2 above). Duanmu, however, takes this step in one sentence on page 26 ('there is no compelling evidence that pre-nasalized stops … exist as single sounds, as I have argued in Duanmu (1994)'), without discussing any piece of evidence, without providing any argument (beyond the fact that his own theory excludes prenasalised consonants), and without informing the reader that there is a plethoric literature, and massive evidence, arguing in favour of a complex segment status for prenasalised consonants. The only literature item quoted is his own 1994 article.

4.4 *When exactly do neighbouring sounds combine into a complex sound?*

The previous sections have shown what a possible complex sound is, and how it comes into being. Let us now consider the question when exactly the potential of a well-formed complex sound is actuated, i.e. when exactly mergeable items merge. Recall that in Duanmu's system the lexicon is made of a linear string of simplex sounds, and that complex sounds are the result of a merger operation. This places an important burden on computation that is usually supported by static inventories. That is, the fact that a language does not have a given segment, say, English the affricate t͡s, becomes an issue of undergeneration in Duanmu's system: if t͡s is a possible complex sound, and if t and s (or s and t since linear order does not matter) may be adjacent in the lexicon, why are they never merged? The author addresses this issue on the occasion of his discussion of tl, dl (30), whose absence in word-initial position in English is as suspect as the absence of t͡s. Duanmu's answer is that languages do not have to use all opportunities to make complex sounds. In other words, the absence of t͡s in English is an accidental gap, as much as the absence of word-initial tl, dl ('English does not need to use all possible complex sounds' (30)).

While in the case of t͡s the statement is unsurprising (in fact the equivalent of the regular inventory-based view), it is incorrect for tl, dl. A trivial and well-known fact about word-initial tl, dl is that they are pervasively absent in many languages, and that equivalent random gaps of, say, #pl, #bl or #kr, #gr, etc. are not on record. Even if nobody has a good explanation why #tl, #dl are missing, it is not the case that their absence is accidental. This is confirmed by the observation that those languages where #tl, #dl do occur also accommodate word-initial clusters that violate sonority sequencing (i.e. #RT, #RR and #TT clusters). For example, this is the case in many Slavic languages (Russian, Czech, Polish, etc., see e.g. Scheer 2007), and in occidental varieties of Modern Arabic (see e.g. Kaye 1990; Scheer 2004: 462–465 on Moroccan Arabic). A reasonable conclusion is that tl, dl are unable to form branching onsets: when they occur, they represent coda–onset sequences, and this is the reason why they cannot exist in word-initial position in languages where initial clusters are restricted to tautosyllabic #TR. This is confirmed by their word-internal behaviour in English: in relevant American

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

varieties, the t of *atlas*, *Atlantic*, etc. is glottalised rather than aspirated (e.g. Kenstowicz 1994: 251). However, glottalisation occurs only in codas.

In his discussion of the absence of t͡s in English (30), Duanmu exposes his general attitude towards overgeneration: there is nothing wrong with it. Complex segments that are well-formed and could occur but do not represent accidental gaps, and there is nothing more to say (this position is also defended by Hale & Reiss 2008). It should be noted, however, that this not only applies to overgeneration within a given system, but also to 'universal' overgeneration: Duanmu's system may well derive well-formed complex sounds that are absent from the record of known natural languages. On page 50, he concludes that '[i]n other words, any stop can combine with any fricative to make an affricate'. Hence affricates as outlandish as p͡x, ɟ͡ʁ, t͡h, k͡θ and the like are held to be just as regular, grammatically speaking, as t͡s and t͡ʃ.

In any event and whatever the analysis of #tl, #dl, Duanmu needs to establish a language-specific list of those mergers producing complex sounds that do, and those that do not, occur. English, for example, must contain a statement that disallows t and s (or s and t) to merge into t͡s. There must also be some equivalent statement in the traditional perspective where the absence of segments is not due to computation, but to lexical inventories. So far so good, but further questions arise. In those cases where the merger of a given pair of sounds into a complex sound is allowed in a particular language, does this merger ALWAYS occur when it can? Or, in other words, is ANY relevant input affected by the computation that creates complex sounds, or is the application of this computation selective? And if so, according to which criteria does it selectively apply?

These are questions that the static lexical alternative does not have to answer. Recall the oft-quoted Polish example from Section 3.2 above that illustrates situations where the same sequence of surface items contrasts according to whether they are mono- or bisegmental: *czy* [t͡ʃɨ] 'whether' vs. *trzy* [tʃɨ] 'three'. On the traditional view, infants acquire the sound inventory of their language during first language acquisition. The inventory is then fixed, stored and disconnected from actual phonological computation that relates input to output forms. In such a system where (in the Polish case) t͡ʃ, t and ʃ are three distinct items of the inventory, the fact that a surface sequence XY is monosegmental in some words, but bisegmental in others, is a simple case of lexical accident, as is the occurrence of any other item of the inventory.

As far as I can see, this is not any different in more recent approaches that attempt to unify the definition of (static) inventories and (dynamic) input–output computation. For example, Steriade (2007: 145f.) argues that 'it is … inaccurate to say that constraints on alphabets … only apply to define the underlying inventory: some of these constraints are persistent (Myers 1991) and prohibit the same feature combinations throughout much or all of the derivation'. The unification of contrast and computation also lies at the

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

heart of Clements (2003, 2009) and Dresher (2009), and the idea is developed in Government Phonology under the heading of Licensing Constraints (Charette & Göksel 1994; Kaye 2001; Scheer 2010b: 2530ff.). In such an environment, the unique set of instructions (constraints) defines what a possible computational event is, AND what a possible item of the static inventory is. But as in the traditional approach, the distribution of items of the inventory over lexical recordings is accidental: in Polish, t͡ʃ, t and ʃ will still be three distinct lexical items whose distribution is unpredictable.

Hence the modern take on inventories, like its predecessor, does not need to make any specific statement regarding the question of what makes XY and X^Y distinct: they are distinct lexically. In Duanmu's system, where complex sounds are banned from the lexicon, however, some additional statement needs to be made in the lexicon in order to differentiate what will appear on the surface as a mono- vs. a bisegmental item. Again, the author does not address this critical issue directly: Duanmu does not mention the literature which provides for a merger mechanism that creates complex segments, but also allows for complex segments to occur in underlying forms. And there is no discussion of well-known cases of the Polish kind: the reader is left guessing what type of lexical diacritic Duanmu could use in order to make surface XY distinct from surface X^Y.

The book only provides an indirect indication: Duanmu is explicit on the fact that the merger computation may be selective. One thing that it depends on is speech rate: 'VNC can be pronounced as three sounds, if one speaks carefully, or it can also be pronounced as two sounds, where VN merges into one sound' (30–31); '[Vʔ]→[$V^{ʔ}$] is found in Shanghai [Chinese], where [Vʔ] can be pronounced as one sound [$V^{ʔ}$] if one speaks fast, or as two sounds if one speaks carefully' (31). This speech-rate–based merger, however, is not applicable to other cases of complex sounds: the monosegmental status of affricates such as t͡s in German is not negotiable. Here the author will have to say that the merger ALWAYS applies, no matter what the speech rate. We are thus left not only with language-specific merger operations, but also with a distinction between those operations that are speech-rate–dependent, and others that are absolute. But even within the latter, this cannot be the full story since, recall, linearity does not matter: an underlying sequence /ts/ should be merged into a complex sound and appear as [t͡s] as much as its mirror image /st/. In the same way, in a language where the transformation of /VN/ into the complex sound Ṽ occurs depending on speech rate, the same surface result should be achieved with an /NV/ input. This is obviously not the case: German has plenty of st clusters, which are pronounced [st], rather than [t͡s], and no French NV sequence ever produces a nasal vowel on the surface.

But one senses from the discussion on page 50 that there is more to the question under which conditions exactly two mergeable items merge.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Duanmu writes that '[o]ne might also ask why *help* is not realised as [hɛp^{l}], where [l] and [p] form a complex sound [p^{l}] in the coda. The answer is that [p] is already supported by "anti-allomorphy" … , and so there is no need to incorporate it into a syllable' (50). Hence merger turns out to be a last resort strategy: it is only when there is no other means to accommodate a sound that it tries to hook onto its neighbour. On the same page, the merging operation also loses against another competitor, sonorant syllabicity: 'one might ask why there is no medial coda [p^{l}] or [k^{l}], which could result from a hypothetical word *neplsa* or *tiklny*. A possible answer is that, because [l] can be syllabic, *neplsa* or *tiklny* can be trisyllabic' (50). Duanmu does not formalise this competition in any way, though.

We may thus note that at least some merger operations are in competition with other options, and that the driving force behind this competition is the accommodation (and ultimately pronunciation) of consonants. All this is just conjecture since the author is not explicit on the existence of competition, or on the way this competition is managed (constraint interaction, etc.; the absence of computational aspects of phonology in the book is discussed at greater length in Section 6 below).

Finally, it appears that merger is a last resort operation only in some cases. Compare the lp in *help* with affricates. The latter always merge no matter what: were their merger also last resort only, a word such as *butcher* would come out with independent t and ʃ where t closes the root syllable: *but.ʃer*. This is because the t can happily be accommodated in the preceding syllable, so there is no 'need' to engage into a merger operation. This is not the case, though: t͡ʃ is an affricate also in *butcher*.

The overall picture thus requires the following variation for the merger operation: (i) statements regarding which items are allowed to merge in a given language; (ii) statements regarding which merger operations are subject to speech rate variation; (iii) statements regarding which merger operations are mandatory (e.g. affricates); and (iv) statements regarding which merger operations are in competition with other syllabic solutions, and how this competition is decided. This is probably all technically workable, but one wonders what the import is of creating this kind of machinery, which is only a direct consequence of the idea that complex sounds are absent from the lexicon. It was mentioned earlier that this ban is entirely unmotivated in the book, that the author stands alone with it, and that arguments in favour of the existence of complex segments in the lexicon that are made in the literature are not discussed. As it stands, Duanmu's system thus adds to the computational burden of the grammar, and must introduce some lexical diacritics in order to describe contrasts of the Polish t͡ʃ vs. tʃ kind. It is hard to believe that this is the right way to go in absence of supporting arguments, and in presence of an alternative that does not need any of the extra machinery.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

4.5 *Stress-based syllabification?*

Duanmu proposes that syllabification is stress-based. In Section 4.5 of the book, he argues against the standard Maximal Onset Principle (MOP), 'according to which consonants between vowels should be syllabified as the onset of the following vowel as far as possible' (56). The reader does not learn, however, how it is decided what is and what is not possible in the standard approach; and – in a book called *Syllable Structure* – sonority sequencing is not mentioned at all.

Duanmu argues that there is no real evidence for the MOP, since English aspiration and flapping can be accounted for by different means (a rare occasion where evidence from phonological processes is considered). The alternative builds on what the author calls the Weight-Stress Principle (WSP): stressed rhymes should be heavy, while unstressed rhymes should be light (59). Syllabification, then, is a function of stress: *át.om* vs. *a.tóm.ic*. On these grounds, p, t, k are aspirated word- and foot-initially (*a.t*^{h}*ómic*), and t, d are flapped when they occur elsewhere (*á*[*ɾ*]*.om*). This mechanism appears to be circular: one wonders how stress is distributed, since the stress-assigning algorithm depends on syllable structure, which, however, does not exist by the time stress is assigned, precisely because it is built on stress.

But beyond this issue, Duanmu seems to really propose stress-based syllabification not as a mere alternative analysis for English, but as a general principle that drives syllabification in all languages, given that this proposal is made in the general part of the book where CVX theory is introduced as such, not in one of the language-specific chapters. In other words, the author intends to refute the MOP by showing that there is no evidence for it IN ENGLISH. He thus leaves unmentioned all those languages where stress does not play any role in segmental phenomena, and where syllabification is a strict function of the MOP. In (Southern varieties of) French for instance, mid vowels are tense if and only if they occur before TRV as in *m*[*e*]*trique* 'metrical', but lax before any other cluster, e.g. *p*[*ɛ*]*rdu* 'lost').[7] Therefore, if Duanmu's analysis of English is correct, the MOP is not operative in this language – but this does not tell us anything about other languages, let alone about the validity of the MOP in syllable theory.

That Duanmu really means stress-based syllabification to be universal is confirmed on page 236, where three core properties of CVX theory are listed: one of them says that '[s]yllabification is determined by the Weight-Stress Principle, according to which the rhyme of a stressed syllable is VX and that of an unstressed syllable is V'. Even if all languages were like English, one wonders what systems with fixed initial (like Czech) or final (like French) stress would look like: heavy syllables could only occur at the left or at the

[7] The detail is exposed in the classic literature on the matter, e.g. Dell (1973: 209ff.), Durand, Slater & Wise (1987), Tranel (1987, 1988).

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

right edge, respectively – a description that does not match the structure of the languages at hand.

5. Anti-allomorphy

5.1 *The potential V analysis*

Let us now turn to the second mechanism used by Duanmu in order to account for consonants that appear not to fit the CVX template: anti-allomorphy. Anti-allomorphy identifies as the following instruction: '[k]eep a morpheme in the same shape regardless of the environment' (47). Duanmu thus entertains a version of what, since the 19th century, has been known as analogy and since then variously, as quoted by Lass (1980: 71), as 'the principle of one-meaning-one-form' (Anttila 1972: 92ff.), 'form–meaning biuniqueness' (Ohlander 1976) and allomorphy minimisation within a paradigm (Kiparsky 1971). More recently, the idea has been present under the headings paradigm uniformity, output–output correspondence (Kenstowicz 1996; Burzio 1996; Benua 1995, 1997 and following) and Optimal Paradigms (McCarthy 2005). As elsewhere, though, Duanmu does not provide discussion of the background: the only reference mentioned is Burzio (1996) (46). While nobody doubts the reality of analogy, it is often hard to tell patterns that are due to analogical activity from those that are the result of regular (synchronic or diachronic) phonological processes (see e.g. the historical overview in Downing, Hall & Raffelsiefen 2005a: 2ff.).

Since neogrammarian times, analogy and phonological processes are viewed as the two basic forces that shape sound patterns by counteracting each other: the former restores regularity by levelling out the irregularity introduced by the latter into paradigms of morphologically related words (e.g. Jespersen 1976 [1922]: 92f., 277f.; Anttila 1972: 94ff.). Viewed from the other end, this means that exceptions to regular sound laws or regular phonological computation may be due to the normalizing action of paradigms, which compares surface forms of the different cells and tends towards invariable morpheme shape. This is what Duanmu calls anti-allomorphy (a term coined by Burzio 1996: 130ff.): analogical action prevents morphemes, in his case roots, as we will see, from appearing in different shapes on the surface.

For example, Duanmu argues that anti-allomorphy is the reason why the *-p* is pronounced in *help* although it is not integrated into any syllable (50, and this is meant literally: not at the underlying level, at no derivational stage and not on the surface either). Indeed, the *-p* in *help* cannot be syllabified in CVX since it is supernumerary with respect to the CVX template, and -lp cannot be a complex sound (it would be pronounced [pl] since linearity does not matter within complex sounds). This is what Duanmu means when he says on several occasions in the book

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

(e.g. page 3) that supernumerary consonants at word edges 'are explained by morphology'.

Since anti-allomorphy is an interaction of two morphologically related forms which produces the non-variation of a given morpheme, there must be some version of the morpheme *help* where the *-p* appears rightfully, i.e. as a duly syllabified item. This condition also precludes the existence of monsters with 24 unsyllabified consonants anywhere, including in the middle of morphemes. That such monsters do not exist, and that supernumerary consonants cannot occur morpheme-internally, is assured by what Duanmu calls the potential V analysis (46). That is, word-final extra consonants can only survive if there are V-initial suffixes that can 'support' them at least in some grammatical forms. Hence the *-p* of *help* can survive by way of anti-allomorphy only because there are V-initial suffixes such as in *help-ing* where the *-p* is properly syllabified in the onset of *-ing*. Were there no V-initial suffixes (or, one supposes, if none of them attached to the specific root *help*), the *-p* could not survive in *help*. The same applies to the left edge of the word: supernumerary initial consonants can only survive if the language has V-final prefixes.

5.2 *Is the syllable inventory of a language predictable from morphology?*

The potential V analysis thus states that supernumerary consonants at word edges are only legal if V-final prefixes and/or V-initial suffixes occur. In Duanmu's approach, though, the correlation also holds true in the other direction: the presence of a certain affix type predicts the occurrence of edge clusters. Hence, 'unlike previous analysis, which attributes the variations [of edge clusters] to syllable parameters, the CVX theory derives the variations from morphology' (51).

The author then provides a typology that compares his morphological with the regular phonological account of cross-linguistic variation of the size of monosyllables. For example, the regular double parameter setting 'no branching codas and no branching onsets' describes a language without extra consonants at either edge; with Duanmu's instruments such a language is characterised as 'no V-final prefixes, no V-initial suffixes'. Another equivalence, describing a language that allows for supernumerary consonants at the left edge, is 'having branching onsets', which in Duanmu's system corresponds to 'having V-final prefixes' (51).

These equivalences do not stand up to fact, and they are also inconsistent with the workings of complex sounds. For example, the last equivalence mentioned above is wrong since languages that allow an extra left-edge consonant – i.e. one that is not syllabified – may well have no branching onsets at all: this is the case in Moroccan Arabic, for example, where both #TR and #RT (as well as #RR and #TT) occur (e.g. *brid* 'to freeze', *rbiT* 'to tie') but TR clusters always behave as heterosyllabic sequences

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

(the preceding vowel shows closed syllable effects, see Kaye 1990; Scheer 2004: 462–465). That is, word-initial #RTV clusters must be bisegmental and heterosyllabic (*r.bɨT*): this conclusion was already reached in Section 4.2 above for Czech and Jiarong. In sum, thus, Moroccan Arabic needs V-final prefixes in order to support supernumerary left-edge consonants, but has no branching onsets – while Duanmu holds that the presence of the former implies the latter.

The same equivalence is also inconsistent with what we know of complex sounds since it treats word-initial branching onsets as clusters whose first member is supernumerary and requires a V-final prefix in order to exist. However, earlier and later in the book TR clusters are said to be complex segments (e.g. in Chapter 8, where all English #TR clusters are analysed like this), which match one single x-slot and hence do not produce any supernumerary consonant. Therefore it is not clear why languages like English where TR clusters are complex segments should be in need of V-final prefixes.

The impact of the preceding needs to be measured in the light of the fact that the typology discussed turns out to be the heart of Duanmu's theory: it is designed to bring home the ambition stated in the subtitle of the book ('the limits of variation'), and the author aims to show that it matches the empirical record better than regular parameters on syllabic constituency. The claim behind this is that phonologists have missed a cross-linguistic generalisation: '[t]he 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' (52). Given the importance that the author grants to this claim, as well as its singularity in linguistic thinking, it may come as a surprise that Duanmu does not make any attempt to support his typological generalisations by cross-linguistic evidence, other empirical material or some case study.

In absence of supporting evidence and given the conflicting record, the reader will have to conclude that there is no such correlation between edge clusters and the occurrence of this or that affix type. That is, the presence of supernumerary consonants (to the extent that one agrees on exactly which consonants are concerned) at word edges is not predictable from the inventory of affixes.

This being said, Duanmu's claim is explicit and excludes certain configurations (still on page 51). Therefore its conditions of falsification can be easily stated. What should not exist if Duanmu is right are languages which have the same set of affixes, but different syllable structure. As far as I can see, though, there are plenty of cases of this kind on record: for example, languages that are closely related genetically and therefore possess the same affix inventory, but follow different syllabic patterns. Polish and Belarusian is such a pair: on the basis of the same set of affixes (in any case V-final prefixes and V-initial suffixes are shared by both languages), the former shows all

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

possible sonority slopes in word-initial clusters (#TR, #RT, #RR, #TT; see Scheer 2007), while the latter is a #TR-only language like English or Spanish, where initial clusters strictly obey sonority sequencing. The presence of V-final suffixes in Belarusian, however, should 'support' stray word-initial consonants.

One may assume that Duanmu's answer is along his general line 'there is nothing wrong with overgeneration' (recall Section 4.4 above): Belarusian grammar does not object to word-initial RT clusters, but for some reason the language does not exploit this potential. For one thing, this argumentation would empty Duanmu's claim of any empirical content since there would be no way anymore to know whether a language should be counted as having branching onsets (or any other syllabic configuration for that matter) or not: a language that is strictly CVCV on the surface could be argued to actually have the potential of forming branching onsets, but for some unknown reason does not actuate this option. In fact all languages could be argued to have the highest syllabic complexity since one never knows whether the non-existing patterns are just accidental gaps.

But in the Belarusian case, even this would not work since there is positive evidence to the end that the language does not tolerate word-initial stray consonants: when a CVC root occurs in zero grade (because a V-initial suffix triggers a vowel-zero alternation on the root vowel, a typical pattern in Slavic), an i-prothesis appears. Compare, for example, *lev* 'lion NOMsg' with *i-lv-a* 'lion GENsg' (Scheer 2009a). If the presence of V-final prefixes guaranteed the viability of the word-initial #lv cluster, the genitive singular could happily be pronounced *lv-a* (as is the case in Polish: *lew – lw-a* 'id.').

Another objection against the existence of a correlation between the phonological properties of affixes and syllable complexity is the simple question how many languages there are where V-final prefixes or V-initial suffixes are absent. This being the critical predictor of Duanmu's system, what kind of cross-linguistic variation do we expect if all languages are the same regarding the conditioning factor? Again, one would expect the different patterns predicted (seven are listed on page 51, but the list ends with 'etc. ') to be illustrated at least by one language, which means that on the affix side Duanmu would have to come up with languages that have 1) V-final prefixes and V-initial suffixes, 2) V-final prefixes but no V-initial suffixes, 3) no V-final prefixes but V-initial suffixes, 4) neither V-final prefixes nor V-initial suffixes. The existence of patterns 2, 3 and 4 is hard to imagine among natural languages.

Finally, one wonders what the predictors 'V-final prefixes' and 'V-initial suffixes' become in languages where syllabification is done across word boundaries (so-called connected speech). Languages where this is notoriously the case include Sardinian (e.g. Bolognesi 1998), Middle Italian dialects (e.g. Marotta 2008, Scheer 2009b), the aforementioned Belarusian (Scheer 2009a), and to a certain extent English. In relevant dialects of the latter,

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

t-flapping also concerns word-final t's, which undergo the process iff the following word is V-initial. Compare, for example, *hit*[*ɾ*] *Ann* with *hit**[*ɾ*] *Mary* (e.g. Nespor & Vogel 1986: 46f., 224f.). We know independently that flapping occurs only in onsets (otherwise the t is glottalised, see Kenstowicz 1994: 251), which means that the t in the external sandhi cases mentioned must sit in the onset of the following word. Now in such a language affixes should not play any role since supernumerary edge consonants can always be supported by a preceding V-final or a following V-initial word. The prediction, then, should be that there is no variation in syllable structure among these languages at all. This is not the case, though. External sandhi languages are not reported to have any particular syllabic typology: their choices regarding syllabic complexity go over the same range of possibilities as those of languages where the domain of syllabification is strictly bound by the word.

5.3 *Law of Finals: If a language has internal codas, it must also have word-final consonants*

The book works with another cross-linguistic generalisation that is controversial and faces a number of counter-examples, but again without inspecting well-known patterns from individual languages. Duanmu claims that there is a correlation between word-final and word-internal rhymes. He calls on Vennemann (1988) and Blevins (2004) (and ultimately on Haugen 1956) in order to motivate the 'Law of Finals (LOF): Word-final rhymes and word-medial rhymes should resemble each other' (55) (there is also a Law of Initials LOI, which does the symmetric labour for the left edge/onsets).[8]

The idea that edge clusters define possible (internal) onsets and possible (internal) codas is standard in the literature. But Duanmu transforms an implication ('if at an edge X exists, then X is a permissible (internal) onset/coda') into an equivalence whereby internal onsets/rhymes also define what kind of cluster occurs at edges. Hence if a language has word-internal codas, it should also have word-final codas (i.e. word-final consonants). By contrast, in case it does not possess word-internal codas, all words should be vowel-final.

A well-known counter-example is Italian, where all words are vowel-final, but word-internal codas regularly occur (*altro* 'other', *parco* 'park', etc.). It is true that word-final consonants occur in words that have recently joined the language (recent loans, acronyms, etc., e.g. *internet* 'Internet', see e.g. Passino 2008, Krämer 2009). These may either be analysed as an instance of a specific grammar (loanword phonology, e.g. Itô & Mester 2004). Alternatively, the pattern 'internal codas but no word-final consonants' is

[8] However, Haugen 1956 is missing from the references section (actually both Haugen 1956a and 1956b, quoted on pages 52–53, are missing), like a number of other items quoted: I have noted Burling (1966) (39), Lamontagne (1993) (50) and Goldsmith (2007) (240).

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

surface-true for earlier stages of Italian, when word-final consonants were actively eliminated, but internal codas survived: Vennemann (1988: 24) states that 'Italian has eliminated (excepting some loan words and truncated forms) all word-final consonants, some by vocalisation as in (31) but most by loss, a few by adding a protective vowel'.

More generally, the correlation between internal codas and word-final consonants has been considered in the literature under the heading of the status of the latter (as a coda or a non-coda). Harris (1994: 162) writes that '[a]ccording to the final-coda view, the typology is exhaustively characterised in terms of a straightforward bifurcation into open-syllable "CV" systems versus closed-syllable "CVC" systems. … The facts, however, contradict this taxonomy. The internal-coda and final-consonant dimensions are actually independent of one another, as a result of which there are four rather than two major types of syllabic system'. Harris then quotes Italian and Telugu (Central Dravidian) as instances of systems with internal codas but no final consonants, and Luo (Western Nilotic, Kenya) as a representative of the reverse pattern (final consonants, but no internal codas).

5.4 *Existing allomorphy: Constraint interaction?*

Finally, we turn to the trivial issue, not addressed in the book, of the bare existence of countless examples of allomorphy. Were anti-allomorphy a rule without fail, morphemes should never appear in different shapes on the surface.[9] The reader senses that Duanmu does not mean to simply say that anti-allomorphy strikes when morphemes do not vary, and is violated when they do. This would void the mechanism of any empirical content: the observation would merely be restated.

Rather, anti-allomorphy looks like a constraint in the sense of OT, but Duanmu does not correlate it with other constraints (as the literature on paradigm uniformity would), or formalise his cross-linguistic typology in terms of constraint interaction.[10] The following section discusses the absence of computational aspects of phonology at greater length.

---

[9] One case from English, the indefinite article *a/an* (*a coffee* vs. *an apple*) is discussed and analysed on pages 64–65, but no mention is made of the fact that according to anti-allomorphy the morpheme should be either always *a* or always *an*.

[10] Except for a short interlude in Chapter 7, where OT-based constraint interaction is adopted and applied to some Standard/Shanghai Chinese data. Elsewhere the reader comes across statements that are called rules: the affix rule (50), the peak rule or the C-rule (both on page 42), and finally there are also laws: the Law of Initials and the Law of Finals (54–56). Of course there is nothing wrong with having a theory that accommodates different kinds of computational instructions (Calabrese 2005 is along these lines), but the reader expects some discussion of Duanmu's view on the matter: are rules, laws and constraints merely interchangeable words, or are they meaningful in the current theoretical landscape? As it stands, the reader is left with an amorphous set of statements that cannot be made sense of in terms of a computational system.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

6. STATICISM

6.1 *Evidence from phonological processes is by and large absent*

A remarkable property of the book is the almost complete absence of phonological processes, something that may be called staticism.[11] Duanmu's main concern is how to accommodate supernumerary consonants. Even though the BEHAVIOUR of these consonants in phonological processes is a classical diagnostic for their syllabic status, the book does not take into consideration this kind of evidence. Even when the issue is pointed out by a reviewer of the book (the lateral of the word *file* is velarised, so it must be a coda, see the following subsection), he rest of the book does not engage with evidence from phonological processes beyond a reaction on the particular issue regarding *file*. Blevins (2010: 290) also mentions this concern in passing: '[i]f *Syllable structure* is to be taken seriously, the same kind of phonological and experimental evidence [including thus phonological processes] needs to be presented for the CVX model'. The discussion below provides some illustration.

For example, staticism drives the section on syllable boundaries (in Chapter 3, 52–64). Duanmu considers how the beginning and the end of syllables can be determined – a classical question in syllable theory. He starts by arguing against speaker intuitions, and then writes 'suppose there is a total lack of speaker intuition, should we conclude that there are no syllable boundaries? I suggest the answer is no' (53). The reader may now expect that the alternative source of evidence that Duanmu proposes are phonological processes. This is not the case. Instead, the aforementioned Law of Initials (LOI) and Law of Finals (LOF) are developed. They are indeed a classical type of evidence as well (though only in their implicational guise, not in Duanmu's equivalence version, see Section 5.3). But significantly, the author does not explore phonological processes, which provide an obvious arbitral award.

Another illustration of staticism is the discussion of English word-internal monomorphemic sequences that look like VXC rhymes. On pages 44–45 and 150–154, relevant items are presented as a problem for CVX theory since they are too big to fit under VX: *symptom* (VNC) and *council* (VVN) illustrate the two patterns. The solution proposed is based on the observation that these items always involve a homorganic NC cluster. Since these clusters are produced by the same place articulator, they do not qualify for complex sound status (see Section 4.3 above). What can be a complex sound, though, is the VN sequence. Therefore Duanmu contends that *ym* in *symptom* and *oun* in

[11] I can only quote two occasions where evidence from phonological processes is used in order to identify syllable structure: on pages 59–63, regarding *t*-flapping and aspiration in English, and on page 72, where fronting and nasalisation of *aa* in Chinese is taken as an indicator for its occurrence in a closed syllable.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

council are complex sounds, that is nasal vowels. This indeed makes the sequences small enough to fit into CVX.

This analysis falls foul of lenition processes in many languages: whenever there is a homorganic NC cluster and a relevant diagnostic for consonantal strength, the C is strong, i.e. shielded against lenition, exactly like other consonants in post-coda position where the coda is not a nasal (e.g. Ségéral & Scheer 2008). Duanmu is explicit on the fact that ALL VNC sequences where N and C are homorganic can be analysed as $\tilde{V}$ ('another case of interest is VNC clusters, where the nasal N and the consonant C are homorganic (sharing the same place of articulation). In this case, VNC can be analyzed as $\tilde{V}$C', 30). The author thus predicts that in English and all other languages where VVN.CV sequences need to be shrunk because they are too big for CVX, the C of the NC cluster is an ordinary intervocalic onset: $\tilde{V}$.CV. This is contrary to fact, though, whenever a lenition process provides a diagnostic for intervocalichood. In English, for example, t-lenition systems show various effects on etymological t's (making them unreleased, a flap, a glottal stop or an s, see Harris 1994: 194ff.). In a number of varieties, t is completely preserved from any damage in post-consonantal position, no matter what the preceding consonant (*chap.ter*, *win.ter*, *coun.ting*, *quar.ter*, *poul.try*, *eas.ter*). This is unexpected for Duanmu, who predicts that t behaves like an intervocalic consonant in *counting* and the like. Now there are also systems on record where t is damaged in intervocalic position AND after nasals (e.g. English spoken in New York: Harris & Kaye 1990). In these varieties lenition also occurs in words like *quarter*, but, contrary to Duanmu's prediction, it is not the nasal vs. non-nasal distinction of the preceding consonant that governs the lenition pattern, but the opposition sonorant vs. non-sonorant. That is, in New York t is flapped in *winter* and *quarter*, but not in *after*, *custard*, *chapter* and *doctor*. This distinction 'post-coda weak after sonorants but strong after obstruents' is recurrent cross-linguistically (Ségéral & Scheer 2008).

6.2 *Supernumerary consonants identified as codas by phonological processes*

Another context in which the absence from the book of reference to phonological processes matters is the unsyllabified status of supernumerary edge consonants. It takes a reviewer of the book manuscript to introduce processing (on page 47 Duanmu reports that he was asked during the reviewing process whether the fact that the lateral in *file* is velarised does not constitute evidence in favour of its coda status). In Duanmu's system the lateral cannot be a coda because the X of CVX is already occupied by the second part of the diphthong. Therefore the lateral ought to be stray (i.e. supported by anti-allomorphy). The author accepts the evidence regarding the darkness of the lateral, but argues that the distribution of clear and dark l does not oppose onset and coda l, but rather laterals that are

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

dominated by onsets vs. rhymes. Since in his system the lateral in *file* cannot be a coda, the author makes it a separate nucleus, which is dominated by a rhyme and therefore velarised. On this analysis, then, the word *file* has a syllabic l and is therefore bisyllabic: $[f[aj]_N]_\sigma\ [[l]_N]_\sigma$.

Beyond the particular case of English dark l (on which more below), the argument concerns a more general pattern: word-final consonants may be subject to coda effects, in which case Duanmu's stray analysis must be wrong. Indeed, processes such as l-darkening, l-vocalisation, r-vocalisation, deaspiration, final devoicing or others of the kind are not known for making any difference between those word-final consonants that, according to Duanmu, are true codas (i.e. which occur after short vowels and therefore instantiate the X of CVX), and those that remain unsyllabified and are supported by a potential V (i.e. which occur after a long vowel or a short vowel plus another consonant and therefore do not fit into CVX).

Let us examine two cases in greater detail, English l-darkening itself and German r-vocalisation. Lavoie & Cohn (1999) study the peculiar properties of *file*-like words in English, i.e. items where a word-final liquid, r or l, is preceded by a long vowel or a diphthong. The intuition of the authors, which coincides with a number of pronouncing dictionaries, is that *file*-type words are longer than one syllable items such as *beat* or *bead*, but not quite as long as two syllable items such as *bottom*. They therefore call them sesquisyllabic words, i.e. words which count as one and a half syllables, and analyse them as threemoraic monosyllables (as opposed to bimoraic monosyllables for the *beat*/*bead* pattern, and a bisyllabic structure for the *bottom*-type items). Lavoie & Cohn support their intuition with evidence from (i) a forced-choice test with six American natives who had to decide how many syllables a written stimulus had, (ii) metrical structure (chant and verse), and (iii) the measurement of duration. The results of the latter show that the vowel of *file*-type words is significantly longer than the vowel of *beat*/*bead*-type words.

On the face of it, this looks like good news for Duanmu, since the extra length of *file*, both measured and perceived, is a fair match of his bisyllabic analysis. Unfortunately, though, this correlation breaks down when we examine exactly which words belong to the *file*-type in Lavoie & Cohn's data: sesquisyllabic words are those which have a word-final lateral preceded by a diphthong (aj, aw, oj like in *pile*, *owl*, *foil*) or a long monophthong, except ɔɔ and what the authors consider to be /o/. Hence *peel*, *pale*, *pearl*, *pool*, *pile*, *owl* and *foil* are extra-long monosyllables, while *Paul* [pɔɔl] and *pole* /pol/ [pəwl] are regular monosyllables. The situation with word-final r is yet different: only preceding diphthongs produce extra length (*pyre*, *hour*, *foyer*); all monophthongs remain regular monosyllables (*peer*, *pear*, *par*, *poor*, *pour*). Hence it is not the case that all l-final words preceded by a long vowel or a diphthong produce the extra length that suits Duanmu's bisyllabic analysis. This is, however, what Duanmu will have to claim since l-darkening, our diagnostic for the syllabic status of word-final l, does not

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

care for the kind of vowel that precedes (long, short, back mid, etc.): all word-final laterals are dark (e.g. Spencer 1996: 214ff.). Therefore, the bisyllabic analysis that Duanmu is forced to recur to sits easily with some long-vowelled items such as *file*, but cannot make sense of others, such as *Paul* and *pole*.

A parallel process is r-vocalisation in German, where r is pronounced as some kind of low schwa [ɐ] in internal and final codas (e.g. Wiese 1996: 252ff.). The r of a stem like *leer* 'empty' surfaces in vocalised form in word-final position (*leer* [leeɐ] 'empty') and before consonants (*leer-te* [leeɐtə] '(s/he) emptied'), but appears as a uvular fricative in intervocalic position (*leer-en* [leeʁ-ən] 'to empty, inf.'). This pattern is entirely insensitive to whether the stem has a long (as in *leer*) or a short vowel (as in *verwirr-en* 'to confuse'). Here as well, Duanmu will have to resort to the rhyme-based analysis of *file*, which in the German case means that the [ɐ] of *leer* [leeɐ] sits in the nucleus of a second syllable, and hence that the item is bisyllabic. Since ɐ is a vowel, its analysis as a nucleus is certainly not outrageous, but one wonders what its status is word-internally, then, as in *leer-te* [leeɐtə]. And in German, the bisyllabic analysis clearly runs against the intuition of speakers, who count unanimously one single syllable for *leer* and the like.

6.3 *Syllabic consonants*

Let us now look at another critical ingredient of the bisyllabic analysis of *file* that Duanmu is forced into: the lateral must be syllabic, i.e. [f[aj]$_{N}$]$_{σ}$ [[l]$_{N}$]$_{σ}$. The trouble is that the syllabicity of consonants – at least on regular assumptions – is connected to the way the items are pronounced: a consonant is not syllabic just because the analyst has decided that it should. It was mentioned above that speakers have intuitions about how many peaks there are in a word. Also, syllabic consonants leave traces in the phonology: they are stress-bearing units, which disqualifies, for example, the lateral in Czech *lhát* 'to lie' and Polish *lgnąć* 'to cling to sth.' from syllabic status. In Czech, stress is initial and would fall on the lateral (but does not: the á is stressed, while in *blbec* 'idiot', for example, stress falls on the lateral); in Polish, stress is penultimate and would also fall on the lateral in *lgnąć*, were it a possible stress-bearing unit (here as well it falls on the ą). Relevant diagnostics for the syllabicity of consonants are discussed in Scheer (2008). The laterals in the Czech and Polish words at hand are thus of a different kind: like their syllabic cousins they are cluster-building and 'supernumerary' from the point of view of ill-formed-looking clusters, but still not syllabic. In the Slavic literature they are called trapped.

Beyond the case of *file*, Duanmu also declares consonants syllabic on other occasions when this suits the analysis. The Tibeto-Burman language Jiarong, studied in Chapter 11, has regular #TR clusters, but also abundantly displays #RT, #RR and #TT sequences. In a first step, Duanmu's account is based

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

on anti-allomorphy: C_1 of any #C_1C_2 is supported by prefix-final vowels and hence remains unsyllabified. Unlike in English, where all #TR clusters are interpreted as complex sounds (159-179), in this language even #TR clusters are analysed as bisegmental (although there is no apparent reason, and no discussion), and the T is thus supported by a potential V as well.

In a second step, clusters bigger than #CC that cannot be accounted for by anti-allomorphy are declared to have a syllabic consonant: this is the case for words and clusters such as *zndai* 'wall', #wrŋ-, #vrd-, etc., which Duanmu analyses as *zn̩dai*, #wr̩ŋ-, #vr̩d-. As it stands, this move certainly saves CVX from Jiarong monster clusters, but there is no evidence that the consonants which Duanmu needs to make syllabic really are. More general information about the occurrence and behaviour of syllabic consonants in Jiarong would be needed, and the application of regular diagnostics (stress, poetry, native speakers' intuitions about the number of syllables) would be required, also regarding the eventual difference between syllabic and trapped consonants.

6.4 *Affix and 'affix-like' sounds are for free*

On pages 49–50, Duanmu is forced into a further radicalisation of the allowance for stray consonants because of words like *saved*, *risked*, *texts*, *sixths*. These are even bigger than VXC where C is already supernumerary: the *-d* of *sav-ed* [sɛjvd], for example, is on top of VXC – the diphthong takes up VX and the v is the extra C. While anti-allomorphy can account for the stem-final -v, it is without effect for the -d since no further V-initial suffix can ever be added. The critical observation is that all such monster clusters involve a morpheme boundary, and that thus morphology has a word to say. In Government Phonology, for example, the conclusion is that each morpheme involved sits in a cycle of its own (cycles are called domains in this theory), and that cycle-final empty nuclei may remain empty (Kaye 1995). Hence *sixths* is made of three cycles and identifies as [[[sɪxø]θø]sø], where three domain-final empty nuclei occur in a row.

Duanmu's solution is also based on this morphology-driven insight, but simply places affix material outside of phonology altogether: the Affix Rule states that '[a]ffix or affix-like sounds can be pronounced, whether they can fit into a syllable or not' (50). In other words, affixes are not concerned by phonology at all and therefore may be stray. As on other occasions, a functional explanation is advanced: if affix material were left unpronounced every time it cannot be syllabified, people would not be able to distinguish affixed from unaffixed forms, and the meaning carried by the affix would be lost (49).

Duanmu's Affix Rule (but not necessarily its functional motivation) expresses the idea known under the heading of the OT constraint Realize-Morpheme (or Realize-Affix), which, in Walker's (1998: 244) formulation, states that '[a] morpheme must have some phonological exponent in the

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

output'. Relevant literature also includes Rose (1997) and Kurisu (2001); the book, however, does not mention the fact that the Affix Rule has a precedent.

But Duanmu goes still farther than that when he talks about perceived affixes and phonetic familiarity (49): since 'German and English use coronal consonants as suffixes, all final coronals are perceived as suffixes' (49, also 214).[12] He therefore introduces a principle according to which '[s]ounds that resemble affixes can be treated as affixes' (49). This is a version of what appears as aggressive suffixation in the literature (but this fact is not mentioned). Zuraw (2002: 396) reports that the term goes back to Hammond (1999b), who 'proposes that English adjectives whose final syllable resembles a suffix are stressed as though they actually were suffixed'.

Finally, according to Duanmu, phonetic familiarity can also explain the presence of extra word-final coronals in monomorphemic items: 'because we are used to hearing coronal suffixes, other final coronals would not sound bad to the ear' (49). The author thus contends that syllabification, at least in part, is a matter of sounding good or bad to the ear. This is probably not a trivial notion, and the reader wonders how it can be determined in practice whether a given sound, or a sequence thereof, is euphonic to the ear.

7. Lexical exhaustiveness and the use of corpora

Language-specific chapters of the book are based to a fair extent on statistical appraisals of the lexicon: How many different syllables could this language have, given its segment inventory? How many does it really have? Where are the missing ones? Are the missing ones systematic or accidental gaps? What is the homophone density? How frequent are sounds, syllables, syllable types, etc.? The statistics at hand are corpus-based, and the use of corpora has an explicit (methodological) ambition: '[w]hile this study aims to offer a general theory of syllable structure, equal emphasis is placed on data description. For each language, quantitative data will be provided from entire lexicons or entire syllable inventories' (5).

There is no doubt that testing theories by running them against exhaustive lexical material is a desirable and recommendable thing to do (e.g. Karttunen 2006). The era of searchable electronic corpora has placed this ambition within the reach of linguists. The book, however, does not always bring

[12] But there are also cases where in order to make his analysis work Duanmu gets rid of substrings which are not anything that one would call an affix by regular standards, and which are not declared perceived affixes either. On page 219 the author gets the German words *Bürste*, *Gerste*, *Fenster*, *Hamster*, *Monster* or *Polster* (and a few others) back to (his) regularity by 'removing the suffix'. This is done without argument or discussion, as if the morphological structure *Bürst-e*, *Gerst-e*, *Fenst-er*, *Hamst-er*, *Monst-er* or *Polst-er* were self-evident (perhaps the CELEX lexicon is the source of the information? – see Section 7 below).

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

home the promise of lexical exhaustiveness. Recall Duanmu's analysis of *symptom* and *council* from Section 6.1 above, whereby the offending VNC and VVN rhymes are shrunk to ṼC. The reader familiar with the literature on superheavy rhymes in English (e.g. Myers 1987; Lamontagne 1993: 147ff.; Harris 1994: 66ff.; Hammond 1999a: 127ff.; Hall 2001), though, is missing items with non-nasal codas such as in *shoulder*, *boulder*, *cauldron*, *poultry*, *smoulder*, *fealty*, *realty*, *holster*, *bolster*, *easter*, *oyster*, *pastry* or *boisterous*.

On pages 149–154, Duanmu analyses English rhymes in non-final position. Starting with the CELEX lexicon that contains 160,595 entries, he first removes compounds, affixed words, acronyms, outlandish scientific terminology that nobody knows, etc. He then isolates those items that still have rhymes bigger than VX: 146 words on his count. After another round of cleaning up (*dachshund* and the like), after reanalysing the x of the prefix *ex-* as an affricate k^s (e.g. *exchange*), and after making *scherzo scher.zo* (rather than *schert.so*), 106 offending items are left. Of these, 99 fall into the VVN (*council*) or VNC (*symptom*) class. The reader is thus left with seven true exceptions where the preceding coda is not a nasal: *arctic*, *dextrose*, *maestro*, *ordnance*, *parsnip*, *poultice* and *seismic* (listed on page 153). These are accounted for with reference to the aforementioned perceived affixes (speakers really treat *arctic* as *arc-tic*, on the model of *drama – drama-tic*).

But where are our *shoulder*, *boulder*, *cauldron*, *poultry*, *smoulder*, *fealty*, *realty*, *holster*, *bolster*, *easter*, *oyster*, *pastry* and *boisterous* (many of which are mentioned in Hall 2001, which Duanmu often quotes)? They must be present in the CELEX database, they qualify for the first search criterion 'rhyme bigger than VX' and thus should be among the output of 146 words, and they are not eliminated by any of the steps that shrink the list of counter-examples to seven. They must thus have been lost underway.

As a result, the author explains that his approach based on lexical exhaustiveness is unlike previous analyses that draw generalisations from fragmentary data – but then for some reason misses out on the offenders that are most commonly quoted in the literature. Corpus-based linguistics is no doubt a good thing to do – but it needs to be done properly, and looking at what the pre–electronic-corpus literature has produced is also recommendable.

The book illustrates yet another risk when annotated corpora are used: the annotations are made by humans, who make theory-sensitive or otherwise motivated decisions when enriching the data. Using annotations like real data without questioning their status may produce strongly biased results that reflect the intentions of the annotators more than actual linguistic fact.

In the chapter on German (209–223), Duanmu sets out to run the universal CVX limitation against the German record, which is reputed to be wild. Regarding the word-internal situation, he therefore wants to know how many words there are exactly which display a rhyme that is larger than what he predicts to be the maximal size, i.e. VCC or VVC. He identifies 303 tokens

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

(in 296 words, page 210), and these numbers are the search result based on the CELEX German lexicon (Baayen, Piepenbrock & Gulikers 1993).

The question is how Duanmu could achieve this result given that rhymes are not anything that ordinary word lists identify. He could indeed search only for rhymes because the CELEX lexicon is in fact syllabified: e.g. *reimen* 'to rhyme' is tagged as [rai][mən] and [CVV][CVC].[13] The reader is informed of this fact, but the only thing that he or she learns about how syllabification was done by the CELEX authors is a hint on page 150 that the Maximal Onset Principle was applied. This, however, concerns the CELEX database for English, and there is no guarantee that the same principles were applied to the German database. Obviously, there are various ways in which strings can be syllabified: a number of decisions need to be made, and syllabification is always according to a particular theory. In the above example *reimen* for instance, somebody has decided that the last item of the first syllable is a vowel rather than an off-glide. And on pages 215–216 the reader discovers incidentally that CELEX identifies amibsyllabic consonants (like the l in *bellen* [bɛlən] 'to bark', which according to the CELEX tags belongs to both syllables of the word). Ambisyllabicity, though, is a controversial analysis to say the least (e.g. Jensen 2000) – but in any case it is an ANALYSIS, not a fact.

These matters are transparent in the book: the work of the CELEX annotators is an unquestioned input to Duanmu's reasoning, uncommented and undiscussed. Their decisions are responsible for how Duanmu's theory fares when it is applied to the data. Or rather, Duanmu's theory is not tested on the grounds of language data – it is tested on the basis of a specific view of syllable structure, and the reader is not informed what this view actually is.

8. CONCLUSION

Although the solutions proposed by Duanmu do not seem to always stand up to the empirical and theoretical challenges that are raised in the book, the issues that are addressed are certainly relevant for phonological theory in general, and in the current development of the field in particular. If only implicitly, Duanmu calls for a rebalancing of sources of phonological explanation in favour of representations. I have argued that while this appears to be necessary after a period of computational fervour, there is no

[13] CELEX also provides phonetic transcriptions (syllabified as well), which, however, appear to be rife with mistakes. Examples on page 214 are *prompt* 'prompt', *Kölsch* 'beer brewed in Cologne, or language spoken in this city', *hübsch* 'pretty' and *Barsch* 'perch (fish) ', all of which CELEX notes with a long vowel where the only possible pronunciation (at least in 'Standard' German) is with a short vowel. Some of these mistranscriptions are critical for Duanmu's demonstration, and the author is forced to appeal to the intuitions of a reviewer (of the book) who confirms that the vowels are short. This is a rather scary window into the empirical content of the CELEX database, which does not appear to be reliable. Another case is found on pages 218 and 200, where *Ingwer* 'ginger' appears with a [ŋgv] cluster. As far as I can see, the only possible pronunciation is *I*[ŋv]*er* without [g].

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

reason to neglect dynamic aspects of phonology. Also, Duanmu's book brings to the fore the idea that syllable structure could be reduced to a sequence of invariant syllabic building blocks. While it may be questioned whether his particular solution is workable, an alternative instantiation of the same idea is available (CVCV) and needs to be evaluated against the same range of empirical, theoretical and scholarly record.

REFERENCES

Anderson, Stephen R. 1976. Nasal consonants and the internal structure of segments. *Language* 52, 326–344.

Anderson, Stephen R. 1985. *Phonology in the Twentieth Century*. Chicago, IL: University of Chicago Press.

Anttila, Raimo. 1972. *An introduction to historical and comparative linguistics*. New York: Macmillan.

Baayen, R. Harald, Richard Piepenbrock & Leon Gulikers. 1993. *The CELEX lexical database* (CD ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Benua, Laura. 1995. Identity effects in morphological truncation. In Jill Beckman, Suzanne Urbanczyk & Laura Walsh Dickey (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers in Linguistics 18), 77–136. Amherst, MA: Graduate Linguistic Student Association.

Benua, Laura. 1997. *Transderivational identity: Phonological relations between words*. Ph.D. dissertation, University of Massachusetts at Amherst.

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Blevins, Juliette. 2010. Review of *Syllable structure: The limits of variation* by San Duanmu, 2009. *Linguistic Typology* 14, 287–291.

Bolognesi, Roberto. 1998. *The phonology of Campidanian Sardinian*. Dordrecht: HIL.

Browman, Catherine P. & Louis M. Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6, 201–251.

Burzio, Luigi. 1996. Surface constraints versus underlying representation. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, vol. 1, 123–141. Salford: European Studies Research Institute (ESRI).

Byrd, Dani. 1996a. A Phase Window Framework for articulatory timing. *Phonology* 13, 139–169.

Byrd, Dani. 1996b. Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24, 209–244.

Calabrese, Andrea. 2005. *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter.

Charette, Monik & Asli Göksel. 1994. Vowel harmony and switching in Turkic languages. *SOAS Working Papers in Linguistics and Phonetics* 4, 31–52. [Also in Henryk Kardela & Bogdan Szymanek (eds.), *A Festschrift for Edmund Gussmann*, 29–56. Lublin: University Press of the Catholic University of Lublin, 1996.]

Clements, George N. 1985. The geometry of phonological features. *Phonology* 2, 225–252.

Clements, George N. 1999. Affricates as noncontoured stops. In Osamu Fujimura, Brian D. Joseph & Bohumil Palek (eds.), *LP '98: Item Order in Language and Speech*, vol. 1, 271–299. Prague: Karolinum Press.

Clements, George N. 2003. Feature economy in sound systems. *Phonology* 20, 287–333.

Clements, George N. 2009. The role of features in phonological inventories. In Eric Raimy & Charles Cairns (eds.), *Contemporary views on architecture and representations in phonological theory*, 19–68. Cambridge, MA: MIT Press.

Clements, George N. & Samuel J. Keyser. 1983. *CV Phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Cyran, Eugeniusz. 2010. *Complexity scales and licensing in phonology*. Berlin: Mouton de Gruyter.
de Lacy, Paul (ed.). 2007. *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press.
Dell, François. 1973. *Les règles et les sons*. Paris: Hermann [2nd edn. in 1985].
Downing, Laura J., T. Alan Hall & Renate Raffelsiefen. 2005a. Introduction: The role of paradigms in phonological theory. In Downing et al. (eds.), 1–16.
Downing, Laura J., T. Alan Hall & Renate Raffelsiefen (eds.). 2005. *Paradigms in phonological theory*. Oxford: Oxford University Press.
Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
Duanmu, San. 1994. Against contour tone units. *Linguistic Inquiry* 25, 555–608.
Durand, Jacques, Catherine Slater & Hilary Wise. 1987. Observations on schwa in Southern French. *Linguistics* 25, 983–1004.
Ewen, Colin. 1982. The internal structure of complex segments. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations: Part 2*, 27–67. Dordrecht: Foris.
Fromkin, Victoria A. 1971. The non-anomalous nature of anomalous utterances. *Language* 47, 27–52.
Fujimura, Osamu & Julie B. Lovins. 1978. Syllables as concatenative phonetic units. In Joan Bybee Hooper & Alan Bell (eds.), *Syllables and segments*, 107–120. Amsterdam: North Holland.
Gregová, Renáta. 2009. Book review: Duanmu 2009, *Syllable structure: The limits of variation*. *SKASE Journal of Theoretical Linguistics* 6, 93–99.
Hale, Mark & Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.
Hall, T. A[lan]. 2001. The distribution of superheavy syllables in Modern English. *Folia Linguistica* 35, 399–442.
Hammond, Michael. 1999a. *The phonology of English: A prosodic optimality-theoretic approach*. Cambridge: Cambridge University Press.
Hammond, Michael. 1999b. English stress and cranberry morphs. Presented at the 73rd Annual Meeting of the LSA, Los Angeles.
Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
Harris, John & Jonathan Kaye. 1990. A tale of two cities: London glottaling and New York City tapping. *The Linguistic Review* 7, 251–274.
Herbert, R. K. 1977. *Language universals, markedness theory, and natural phonetic processes: The interaction of nasal and oral consonants*. Ph.D. dissertation, The Ohio State University.
Hirst, Daniel. 1985. Linearisation and the Single-Segment Hypothesis. In Jacqueline Guéron, Hans Obenauer & Jean-Yves Pollock (eds.), *Grammatical representation*, 87–100. Dordrecht: Foris.
Hoard, James E. 1967. *On the foundations of phonological theory*. Ph.D. dissertation, University of Washington.
Hoard, James E. 1978. Syllabication in Northwest Indian languages, with remarks on the nature of syllabic stops and affricates. In Alan Bell & Joan Bybee Hooper (eds.), *Syllables and segments*, 59–72. Amsterdam: North Holland.
Hualde, José. 1988. Affricates are not contour segments. *The Seventh West Coast Conference on Formal Linguistics* (WCCFL 7), 77–89.
Hualde, José. 1991. *Basque phonology*. London: Routledge.
Itô, Junko & Armin Mester. 2004. The phonological lexicon. In John J. McCarthy (ed.), *Optimality Theory in phonology*, 552–568. Oxford: Blackwell.
Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.
Jensen, John T. 2000. Against ambisyllabicity. *Phonology* 17, 187–235.
Jespersen, Otto. 1976 [1922]. *Nature, évolution et origines du langage*. Paris: Payot.
Karttunen, Lauri. 2006. The insufficiency of paper-and-pencil linguistics: The case of Finnish prosody. In Miriam Butt, Mary Dalrymple & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 287–300. Stanford, CA: CSLI Publications. [ROA #818]

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Kaye, Jonathan Derek. 1985. On the syllable structure of certain West African languages. In Didier Goyvaerts (ed.), *African linguistics: Essays in memory of M. W. K. Semikenke*, 285–308. Amsterdam: John Benjamins.
Kaye, Jonathan. 1990. Government in phonology: The case of Moroccan Arabic. *The Linguistic Review* 6, 131–159.
Kaye, Jonathan. 1995. Derivations and interfaces. In Jacques Durand & Francis Katamba (eds.), *Frontiers of phonology*, 289–332. London & New York: Longman. [Also in *SOAS Working Papers in Linguistics and Phonetics* 3, 90–126 (1993).]
Kaye, Jonathan. 2001. Working with licensing constraints. In Katarzyna Dziubalska-Kołaczyk (ed.), *Constraints and preferences*, 251–268. Berlin & New York: Mouton de Gruyter.
Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7, 193–231.
Kehrein, Wolfgang. 2002. *Phonological representation and phonetic phrasing: Affricates and laryngeals*. Tübingen: Niemeyer.
Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Oxford: Blackwell.
Kenstowicz, Michael. 1996. Base-identity and uniform exponence: Alternatives to cyclicity. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, vol. 1, 363–393. Salford: European Studies Research Institute (ESRI).
Kim, Hyunsoon. 1997. *The phonological representation of affricates: Evidence from Korean and other languages*. Ph.D. dissertation, Cornell University.
Kim, Hyunsoon. 2001. A phonetically based account of phonological stop assibilation. *Phonology* 18, 81–108.
Kiparsky, Paul. 1971. Historical linguistics. In William Orr Dingwall (ed.), *A survey of linguistic science*, 576–649. College Park, MD: University of Maryland Linguistics Program.
Krämer, Martin. 2009. *The phonology of Italian*. Oxford: Oxford University Press.
Kurisu, Kazutaka. 2001. *The phonology of morpheme realization*. Ph.D. dissertation, University of California, Santa Cruz.
La Charité, Darlène. 1993. *The internal structure of affricates*. Ph.D. dissertation, University of Ottawa.
Lamontagne, Gregory A. 1993. *Syllabification and consonant cooccurrence conditions*. Ph.D. dissertation, University of Massachusetts.
Lass, Roger. 1980. *On explaining language change*. Cambridge: Cambridge University Press.
Lavoie, Lisa & Abigail C. Cohn. 1999. Sequisyllables on English: The structure of vowel-liquid syllables. *The 14th International Congress of Phonetic Sciences* (ICPhS 14), 109–112.
Levin, Juliette. 1985. *A metrical theory of syllabicity*. Ph.D. dissertation, MIT.
Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language & Linguistic Theory* 8, 375–425.
Lowenstamm, Jean. 1996. CV as the only syllable type. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, vol. 2, 419–441. Salford: European Studies Research Institute (ESRI).
Lowenstamm, Jean. 2003. Remarks on mutae cum liquida and branching onsets. In Stefan Ploch (ed.), *Living on the edge: 28 papers in honour of Jonathan Kaye*, 339–363. Berlin & New York: Mouton de Gruyter.
Maddieson, Ian. 1989. Prenasalized stops and speech timing. *Journal of the International Phonetic Association* 19, 57–66.
Marotta, Giovanna. 2008. Lenition in Tuscan Italian (Gorgia Toscana). In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition*, 235–271. Berlin: de Gruyter.
Martinet, André. 1939. Un ou deux phonèmes? *Acta Linguistica* 1, 94–103.
McCarthy, John J. 2005. Optimal paradigms. In Downing et al. (eds.), 170–210.
McCarthy, John J. & Alan Prince 1996. Prosodic morphology 1986. Ms., University of Massachusetts, Rutgers University.
Myers, Scott. 1987. Vowel shortening in English. *Natural Language & Linguistic Theory* 5, 485–518.
Myers, Scott. 1991. Persistent rules. *Linguistic Inquiry* 22, 315–344.
Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
Ohlander, Sölve. 1976. *Phonology, meaning, morphology*. Göteborg: Acta Universitatis Gothoburgensis.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Passino, Diana. 2008. *Aspects of consonantal lengthening in Italian*. Padova: Unipress.
Rennison, John R. 1998. Contour segments without subsegmental structures. In Eugeniusz Cyran (ed.), *Structure and interpretation: Studies in phonology*, 227–245. Lublin: Folium.
Rose, Sharon. 1997. *Theoretical issues in comparative Ethio-Semitic phonology and morphology*. Ph.D. dissertation, McGill University.
Rosenthall, Sam. 1992. Prenasalized stops and feature geometry. In Wolfgang Dressler, Hans Luschützky, Oskar Pfeiffer & John Rennison (eds.), *Phonologica 1988: The 6th International Phonology Meeting*, 249–258. Cambridge: Cambridge University Press.
Rubach, Jerzy. 1994. Affricates as strident stops in Polish. *Linguistic Inquiry* 25, 119–143.
Sagey, Elizabeth Caroline. 1986. *The representation of features and relations in nonlinear phonology*. Ph.D. dissertation, MIT.
Scheer, Tobias. 2004. *A lateral theory of phonology*, vol. 1: *What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.
Scheer, Tobias. 2007. On the status of word-initial clusters in Slavic (and elsewhere). In Richard Compton, Magdalena Goledzinowska & Ulyana Savchenko (eds.), *Annual Workshop on Formal Approaches to Slavic Linguistics: The Toronto Meeting 2006*, 346–364. Ann Arbor, MI: Michigan Slavic Publications.
Scheer, Tobias. 2008. Syllabic and trapped consonants in (Western) Slavic: The same but yet different. In Gerhild Zybatow, Luka Szucsich, Uwe Junghanns & Roland Meyer (eds.), *Formal description of Slavic languages: The Fifth Conference, Leipzig 2003*, 149–167. Frankfurt am Main: Lang.
Scheer, Tobias. 2009a. Representational and procedural sandhi killers: Diagnostics, distribution, behaviour. In Mojmír Dočekal & Markéta Ziková (eds.), *Czech in formal grammar*, 155–174. München: Lincom.
Scheer, Tobias. 2009b. External sandhi: What the initial CV is initial of. *Studi e Saggi Linguistici* 47, 43–82.
Scheer, Tobias. 2010a. What OT is, and what it is not. Review article on *The Cambridge Handbook of Phonology*, edited by Paul de Lacy, 2007. *Journal of Linguistics* 46.1, 193–218.
Scheer, Tobias. 2010b. How to marry (structuralist) contrast and (generative) processing. Review of *The Contrastive Hierarchy in Phonology* by Wolfgang Dresher, 2009. *Lingua* 120, 2522–2534.
Scheer, Tobias. 2011. Aspects of the development of generative phonology. In Bert Botma, Nancy Kula & Kuniya Nasukawa (eds.), *The Continuum companion to phonology*, 397–446. New York: Continuum.
Ségéral, Philippe & Tobias Scheer. 2008. Positional factors in lenition and fortition. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition*, 131–172. Berlin: Mouton de Gruyter.
Spencer, Andrew. 1996. *Phonology*. Oxford: Blackwell.
Steriade, Donca. 1993. Closure, release, and nasal contours. In Marie K. Huffman & Rena A. Krakow (eds.), *Nasals, nasalization, and the velum*, 401–470. Orlando, FL: Academic Press.
Steriade, Donca. 1994. Complex onsets as single segments: The Mazateco pattern. In Jennifer Cole & Charles Kisseberth (eds.), *Perspectives in phonology*, 203–291. Stanford, CA: CSLI Publications.
Steriade, Donca. 2007. Contrast. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, 139–157. Cambridge: Cambridge University Press.
Szigetvári, Péter & Tobias Scheer. 2005. Unified representations for the syllable and stress. *Phonology* 22, 37–75.
Tranel, Bernard. 1987. Floating schwas and closed syllable adjustment in French. In Wolfgang Dressler, Hans Luschützky, Oskar Pfeiffer & John Rennison (eds.), *Phonologica 1984*, 311–317. London: Cambridge University Press.
Tranel, Bernard. 1988. A propos de l'ajustement de e en français. In Paul Verluyten (ed.), *La phonologie du schwa français*, 89–131. Amsterdam & Philadelphia, PA: John Benjamins.
van de Weijer, Jeroen. 1992. Basque affricates and the Manner–Place dependency. *Lingua* 88, 129–147.
van de Weijer, Jeroen. 1993. The Manner–Place dependency in complex segments. *Linguistics* 31, 87–110.
van de Weijer, Jeroen. 1996. *Segmental structure and complex segments*. Tübingen: Niemeyer.

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press

Vennemann, Theo. 1988. *Preference laws for syllable structure and the explanation of sound change*. Berlin: Mouton de Gruyter.
Walker, Rachel. 1998. *Nasalization, neutral segments and opacity effects*. Ph.D. dissertation, University of California, Santa Cruz.
Wiese, Richard. 1996. *The phonology of German*. Oxford: Oxford University Press.
Zuraw, Kie. 2002. Aggressive reduplication. *Phonology* 19, 395–439.

Author's address: *Université de Nice – Sophia Antipolis, Campus Saint-Jean d'Angély, Laboratoire Bases, Corpus, Langage (BCL), 24, avenue des Diables Bleus, F-06000 Nice, France*
scheer@unice.fr

https://doi.org/10.1017/S0022226712000205 Published online by Cambridge University Press