To appear. In Kuniya Nasukawa (ed.), *Morpheme-internal Recursion in Phonology* [SGG series]. Berlin/Boston: Mouton de Gruyter, 21–32.
This work is the result of a collaborative research effort carried out by the two authors, both of whom contributed equally to the published paper.

Edoardo Cavirani and Marc van Oostendorp

A theory of the theory of vowels

1 Introduction

One line of research in theories of phonological primitives, such as features or elements, is to reduce the number of such primitives. Feature geometry theories can be seen as a way to do this: a segment is a tree that consists of a treelet (a subtree) of place features, a treelet of aperture features, etc. Especially within Element Theory, there has been a tendency to reduce the number of representational primitives also in other ways. A radical example of this is so-called Radical CV Phonology by Van der Hulst (1988, 1994, 1996, 2015), which claims that there are only two such primitives, called C and V, which can be interpreted differently in different parts of the tree.

Would it be possible to reduce the number of primitives even further? And what would the ultimate reduction be? Would there be a possibility of giving a representation for segments that would not include *any* elements? We know that graph theory, in particular the theory of trees, that already plays such an important role in grammatical theory, is a rather strong mathematical tool – could we not build a theory of vowels on it? This would imply that we have a theory in which the original elements |A, I, U| themselves have internal structure, and such structure would only be treelets. We think we can, and the following can be seen as a complete definition of vowel representations in natural language motivated by the *desiderata* above.

It is the purpose of this paper to sketch what such a theory would look like. By doing so, we are basically building a metatheory of Element Theory. Our 'data' are most vocalic systems that can be successfully derived from the elements |A, U, I| and combinations thereof. We try to derive such systems making no use of elements. The resulting theory would therefore have the same consequences as 'normal' element theories, but it would give an answer to questions such as why there are three elements and why these elements each have their own special properties.

2 Vowels without features of elements. A proposal

We first define the vowel recursively, assuming that we have treelets, i.e. structures that consist of a mother node and one or two daughters, where the daughters can themselves be further treelets:

(1) a. The empty node (also written as ∅) is a vowel, i.e. a treelet.
b. If σ, τ are treelets, the structure in which σ dominates τ is also a treelet.
c. If σ, τ, υ are treelets, and {τ, υ} are balanced, the structure in which σ dominates τ and υ is also a treelet.

The definition in (1c) relies on a definition of 'balance' that is familiar from search algorithms, and to which we will return below (we will ignore it for now). The definition in (1) gives us in principle an infinite number of treelets, of which the following are some simple examples:

(2) a. ∅ (1a)

b.
```
 ∅
 |      (1b)
 ∅
```

ci.
```
  ∅
 / \    (1c, with σ, τ and υ empty)
∅   ∅
```

cii.
```
  ∅
 / \    (1c, with σ, τ empty and υ = (2b))
∅   |
    ∅
```

```
ciii.  Ø
      / \        (1c, with σ empty and τ, υ = (2b))
      |  |
     Ø  Ø
```

The Ø signs in these treelets are *not* labels, and in fact have no special status. We put them here purely for reasons of clarity; leaving them out would give the same mathematical structures. Also, linear order is irrelevant, so that (cii) could also be represented as:

```
(3)   Ø
     / \
     | Ø
    Ø
```

Further treelets can be formed by replacing daughter treelets in any of these representations by another treelet. For instance, we can take (2ciii) and replace one of its daughters by (2ci); this is because the definition in (1) is recursive:

```
(4)   Ø          Ø          Ø
     / \    +   / \    >   / \
     |  |       Ø Ø        | / \
    Ø   Ø                  ØØ Ø
```

In principle, since the original definitions in (1) are recursive, we can go on doing this forever and generate an infinite number of potential vowels. The treelets in (2) have a special status: they are as it were the most primitive structures that exist in the theory, as they are the only ones which can be formed by applying the definitions in (1) at most once. The result of the operation in (4) is more complicated.

There are thus 5 'primitive' treelets. We propose that each of these represents one of the primitives of Element Theory:

```
(5)   Ø        corresponds to an empty x slot

      Ø
      |        corresponds to |@|
      Ø

      Ø
     / \       corresponds to |U|
    Ø   Ø

      Ø
     / \       corresponds to |I|
    Ø   |
        Ø

      Ø
     / \       corresponds to |A|
     |  |
    Ø   Ø
```

Logic teaches us that, since linear order is irrelevant, there can be only three binary treelets, corresponding exactly to the typical number of 'normal' elements, and one unary treelet for the schwa interpretation (in section 2.2, we will go into the question why we think each individual treelet has the interpretation we have provided here).

By assuming that elements are primitive treelets, we have thus derived their number, and by assuming that treelets are recursive, we have also derived another property of Element Theory, viz. that there is no substantive difference between and an element and a combination of elements: they are all (pronounceable) vowels.

In other words, when we combine treelets (per (1ciii)), the resulting tree is again a representation of a vowel, with an interpretation that is familiar from Element Theory, or to be more precise, of Particle Phonology (Schane 1984), because there is no headedness in our representations, and the same treelet can occur more than once. This is true in particular for the |A| treelet. E.g. in (4) above we see |A| combined with |U|, so that the resulting tree is a representation for /o/, and the following would be a representation for /ɔ/:

```
(6)   Ø           Ø           Ø
     / \    +    / \    >    / \
     | |         | /\        |  / \
     Ø Ø         Ø ØØ        Ø  |  /\
                                Ø Ø  Ø
     |a|         /o/          /ɔ/
```

One does not have to have the 'derivational history' of a treelet (as in (7b)) to understand its phonetics; by reading a tree one layer (of (7a)) at a time, one also gets the desired interpretation:

```
(7)  a.   Ø                  b.
         / \                          |A|
        |  / \                      @     |A|
        Ø |  / \                        @    |U|
          Ø Ø  Ø
```

The top layer is a binary treelet with two non-empty daughters; this is interpreted as an |A| element. The second layer consists of a unary set (a schwa) plus another binary treelet with two non-empty daughters, so this is again an |A| element, of which again one of the daughters is a schwa; the other is a treelet with two empty daughters, which is an |U|. This representation can thus be interpreted as the element combination |A|.@.|A|.@.|U|, or, given that @ is an element that does not add any specific value (other than background noise; Harris 1994) as |A|.|A|.|U|, which is an /ɔ/ in Particle Phonology.

2.1 Balance and the limits of recursion

The |A| treelet allows for recursion in this way – limited, we propose, only by extralinguistic factors such as the ability of humans to distinguish vowel heights from each other articulatorily or acoustically. Note that the |U| treelet does not allow any kind of further embedding, because each of its daughters should be empty by definition (if they are not, the treelet is simply not interpreted as |U|). |I| does allow for embedding of |U|, giving us /y/:

```
(8)  a.   Ø                  b.
         / \                     |I|
        Ø  / \                     |U|
           Ø Ø
```

All front rounded vowels will have this treelet as part of their representation. Other types embeddings are universally not allowed, because of what we call the 'balance property' of phonological representations, which we have hitherto not discussed. This property is easily defined informally:

(9) The number of embeddings N(T) in a treelet T, is the number of steps it takes to go from the root of T to the most deeply embedded leaf.

(10) A binary tree {A, B} is **balanced**, if $N(A) – N(B) \leq |1|$ (Adelson-Velsky and Landis 1962)

In other words, the two daughters of a node should have a similar amount of structure: one can be at most one level deeper than the other. If we embed something else than a |@| treelet or a |U| treelet into an |I| treelet, we get a tree that is not balanced. This for instance would be the result of putting an |I| inside another |I|:

```
(11)  a.    Ø                     b.
           / \                        |I|
         Ø / \                          |I|
           Ø  |                           @
              Ø
```

The first daughter of the top node has a number of embedding of 0 (because it is an empty set), the second daughter (the embedded |I|) has a number of embedding of 2, so the difference between the two nodes becomes too big, and there is no balance. The embedding of |U| in |I| is balanced, however (see (8)): the first daughter of |I| still has an embedding of 0, but the second one (the |U|) has an embedding of 1. This difference is within the limits put forward by (10).

Note that this implies that |A| can embed twice into itself, because each daughter is non-empty, and has therefore a number of embedding of at least 1. The representation of /ɔ/ above (7) is therefore balanced. We cannot embed three |A|'s into each other in this way, however, as then the 'simple' daughter of the top node (@) becomes too simple (note that this also derives the (nearly absolute; Crothers 1978) maximal 4-degrees height of vowel systems). We can do more embedding, but in that case, each of the daughters needs more internal structure.

Note that we have tacitly assumed so far that schwa formation is not recursive: we cannot embed a single treelet into another single treelet. The following are not feasible representations:

```
(12)  a.   Ø                b.   Ø
           |                     |
           Ø                     Ø
           |                     |
           Ø                     Ø
                                 |
                                 Ø
```

We assume that the reason for this is that such structure is redundant: a schwa embedded in a schwa would still be phonetically interpreted as a schwa. The second structure in (12) would still be interpreted as @.@.@=@ For this reason, schwa embedding is also not allowed to 'balance' treelets that would otherwise be unbalanced, i.e. we cannot 'save' a recursive |A| by making one arm into a 'big' schwa along the lines of (12).

2.2 Some examples of vowel sets

For the sake of saving space on the page or on your computer screen, we will from now on write our treelets in a kind constituent notation, so that schwa is {Ø}, |U| is [Ø,Ø], etc. Using the kinds of representations just outlined, we can now define a number of well-known vowel sets. For instance, a three-vowel set [u i a] has the following elements:

(13) a. [Ø,Ø] (|U|=[u])
b. [Ø,[Ø]] (|I|=[i])
c. [[Ø],[Ø]] (|A|=[a])

We can call (13a) the |U| treelet, (13b) the |I| treelet and (13c) the |A| treelet. The language does not have schwa, hence no unary treelets. Such a vowel inventory can be described in the following way:

(14) a. All vowels are binary treelets.
b. The treelets that are daughters of vowels are at most unary.

Taken together, (1) and (14) give a precise definition of the language. We assume that (1) is universal: it just defines what it means to be a vowel. (14b) filters out those treelets that are universally available, but not as vowels in this language. (14) is therefore what needs to be acquired by a language learner. This could be a matter of parameter setting. In a four vowel language that also includes schwa, the requirement in (14a) is replaced by (14a)’:

(14) a’. All vowels are monovalent or binary treelts.

If we instead cancel (14b) and add the assumption (14b’) instead, we get a six vowel set (assuming that an |A| set that dominates another |A| set is still also an [a] because of redundancy):

(14) b.’ Vowel sets have an embedding depth of at most 2.

(15) [∅,∅]	(|U|=[u])
[∅,[∅]]	(|I|=[i])
[[∅],[∅]]	(|A|=[a])
[∅, [∅, ∅]]	(|I.U|=[y])
[[∅],[∅,∅]]	(|A.U|=[o])
[[∅],[∅,[∅]]]	(|A.I|=[e])
[[∅],[[∅],[∅]]]	(|A.A|=[a])

Disallowing the front rounded vowels involves adding an extra requirement:

(16) Binary nodes cannot be sisters to empty nodes.

Extending the embedding depth to level 3 instead of level 2, we of course also extend the vowel set even further, viz. to a 5-vowel set:

(17) [∅,∅]	(|U|=[u])
[∅,[∅]]	(|I|=[i])
[[∅],[∅]]	(|A|=[a])
[[∅],[∅,∅]]	(|A.U|=[o])
[[∅],[∅,[∅]]]	(|A.I|=[e])
[[∅],[[∅],[∅]]]	(|A.A|=[a])
[[∅],[[∅],[∅,∅]]	(|A.A.U|=[ɔ])
[[∅],[[∅],[∅,[∅]]]]	(|A.A.I|=[ɛ])
[[∅],[[∅],[[∅],[∅]]]]	(|A.A.A|=[a])

We can understand lowering processes by embedding an |A| vowel in another vowel. In other words, the calculus for at least the most common vowel inventory types can be described with a small set of possible restrictions on sets. Notice that the view on vowel structure which we thus get is not incompatible with autosegmental views of frontness or roundness harmony: these sets can behave as autosegmental elements and spread. Height harmony would need to have a different representation, on the other hand, but the theory is not different in this respect from element theory, as expected.

2.3 Extensions to larger segmental inventories

Obviously, segmental inventories do not just consist of vowels – and even within vocalic phonology there are many distinctions we have not made yet – we will need representations for nasality, for tone, and many other distinctions. Space does not permit to go into these details, but note that we can make use of the kinds of ideas developed in other Element-based frameworks (such as RcvP; see below) in which the same elements can have different phonetic interpretations, depending on their position in the tree. Making segmental trees bigger (having ever more recursion) will expand our space of possibilities.

3 Substance reduction and set theory: some precedents

3.1 Early precursors

Ours is not the first proposal for reducing the number of representational primitives. Several proposals have been put forward with a similar aim. Since our work represents a rather extreme move along this line of research, we think it is important to recapitulate the most relevant stages of this research line. More importantly, this will also give us the room to stress the differences and similarities between our proposal and the preceding ones.

One of the first of this type of measures to reduce the number of representational primitives is Feature Geometry (Clements 1985), which (implicitly and informally) applies the notion of set to the unordered bundles of features of Chomsky and Halle (1968). As a result, segments become sets of subsets of features, which are formally conceived of as organized in 'groups' headed by nodes in a (segmental) tree[1]. Crucially, the geometric restructuring of the featural content of segments allows for generalizations which target subsets, i.e. representational nodes[2]. As a matter of fact, this framework reduces the computational components of phonology (e.g. both the structural description and the structural change of a given rule can now just refer to the relevant parent node), rather than the representational one. Even if representations are still as rich as they were before, though, with Feature Geometry, trees enter the subsegmental scene.

As a matter of fact, trees have already been on the marketplace for phonological theories since a few years (see section 3.5 below). Indeed, assuming the Structural Analogy Hypothesis, whereby both morphosyntactic and phonological structures are represented as dependency relations holding between representational primitives, Anderson and Jones (1974) developed Dependency Phonology (henceforth DP; see also Anderson 1985, 1992; Anderson and Ewen 1987, van der Hulst 2006 and van der Hulst 2011). Within such a model, the organization of features essentially parallels the one proposed by Clements (1985), with major nodes corresponding to laryngeal, manner and place categories. Differently from Feature Geometry, the representational primitives co-occurring under the relevant nodes are arranged according to a variable head-dependent schema: given two features α and β, the relationship they enter into can be either α- or β-headed, each corresponding to a (potentially contrastive) phonological expression. As a matter of fact, dependency relations are suggested to hold also between nodes and sub-nodes. However, no restrictive theory constraining the various combinatorial possibilities has been developed, resulting in overgeneration; see van der Hulst (2006) and (2011) for a brief discussion.

One difference between Feature Geometry and DP concerns the representational primitives, which are binary in the former case and unary in the latter. Furthermore, the primitives proposed by DP are "(in an Aristotelian sense) 'substances' in themselves rather than properties of substances'. Whereas mainstream binary features are arguably properties of segments, DP-primes are segments themselves. Indeed, such primes can occur independently as fully pronounceable phonological segments" (van der Hulst 2006: 455). Traditionally, these primes have been referred to as components, their primary phonetic interpretation being acoustic (e.g. "|V|, a component which can be defined as 'relatively periodic', and |C|, a component of 'periodic energy reduction'" (Anderson and Ewen 1987: 151)).

Many DP proposals were further elaborated within e.g. Radical CV Phonology (henceforth RcvP; van der Hulst 1988, 1994, 1996, 2015) and Government Phonology (henceforth GP; Kaye *et al* 1990; Charette 1991; Lowenstamm 1996 and Scheer 2004). Both RcvP and GP maintain a similar conception of primes, which are unary, 'substantial' and combinable in head-dependent structures. RcvP and GP, though, attempt to solve the overgeneration problem DP suffered due to the lack of a constrained theory of primes (and their combinatorial possibilities). These frameworks proposed two different solutions, which are briefly described in what follows.

1 According to Clements (1985: 230), this embodies the view according to which "the varying degrees of independence among phonetic features can be expressed by a hierarchical grouping such that higher-branching categories tend to be more independent than low-branching categories. More exactly, the relative independence of any two features of feature classes is correlated with the number of nodes that separate them".

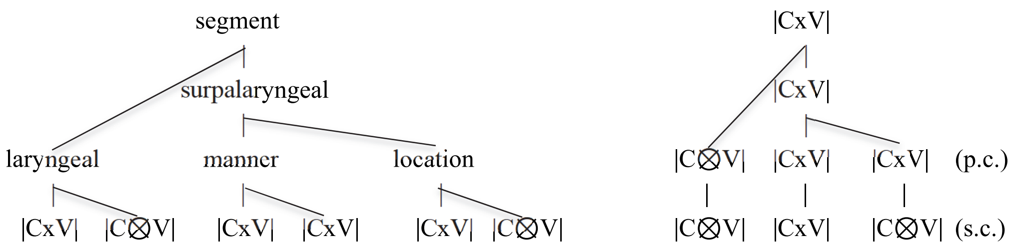
2 More recently, Bale et al. (submitted: 1) resort to set theory in a more explicit fashion: "taking […] feature bundles to be sets [and natural classes sets of sets] allows us to apply ideas from set theory to phonology". This allows them to propose the reconceptualization of a fully underspecified segment as empty set, which, in turn, "can be used to define a natural class over all segments".

3.2 RcvP

In order to limit the generative power of the system developed within DP, RcvP capitalizes on a suggestion already present in Anderson and Ewen (1987), according to which a given component can occur under different nodes of the segmental tree. This is the case, for instance, for the |i| and |u| components, which are interpreted as high and low tone, respectively, when occurring under the tonological node (Anderson and Ewen 1987: 273)[3]. The possibility for the same component to occur in various structural positions, in turn, allows for the formalization of similarities (same component) and differences (different structural position) among (the phonetic interpretation of) segments. Together with the head-dependent asymmetry DP shares with Feature Geometry, this possibility allows for a further reduction of the number of components: a contrast previously formalized by e.g. two features can now be conveyed by one and the same component occurring in a head or dependent guise, or in different structural positions. For instance, the |V| component can translate [sonorant] and [voice] depending on its head *vs* dependent status, or it can identify sonorants, vowels, [low] and [open place] depending on its structural position.

RcvP exploits these possibilities to their maximal extent by constraining the typology of structures to head-dependent configurations of just two primes: |C| and |V| (against the |C|, |V|, |O| |G|, |K|, |i|, |u|, |a|, |@|, |□|, |l|, |t|, |d|, |r|, |L| and |n| of DP), which are organized in an arboreal structure such as the one in (18), where '|CxV|' means that |C| and |V| can combine, '|C⨂V|' that they cannot, and DP gestural labeled nodes (on the left; Clements 1985) are "defined in purely structural terms" (on the right), 'p.c.' and 's.c.' indicating the primary and secondary component, respectively; van der Hulst 2017):

(18) RcvP translation of DP segmental tree



Notice that, even though RcvP, as DP, assumes that components have a default (acoustic) phonetic interpretation, |C| and |V| are assigned specific interpretations depending on their a) syllabic position (onset head *vs* onset dependent *vs* rhyme head *vs* rhyme dependent), b) class (manner *vs* location *vs* laryngeal), c) component status (primary *vs* secondary) and d) element status (head *vs* dependent). RcvP therefore comes as close to completely reducing the role of differences between elements as one can get without really abandoning the whole concept of elements completely. It does not seem to make sense, for instance, to develop a model with only *one* primitive. One can then only differentiate different structures by counting; and the same can be obtained by having no primitives at all.

3.3 GP and GP 2.0

As mentioned above, DP-style representational primitives are kept in both RcvP and GP, inasmuch as they all resort to unary primes which are acoustically grounded and combinable in head-dependent structure. In the GP literature, these primes are known as elements. Clearly rooted in DP (and Particle Phonology; Schane 1984), elements are introduced by Kaye et al (1985) and further developed along various directions, which differ in the element number and/or in the way elements can combine (see Backley 2012 for an overview and a brief discussion of the variants on the market).

In its standard form (Backley 2011), there are six elements, which are extensionally equivalent to the objects defined within RcvP by means of elements and (unlabeled) gestural nodes:

[3] Interestingly, Anderson & Ewan (1987: 215) argues for the identity of |a| and |V|. As we will see below, this alleged identity is in line with our proposal, as well as with those proposed e.g. by Rennison (1998) within the GP camp.

(19) GP *vs* RcvP (van der Hulst 2017)

GP elements	
|ʔ|	|A|
|I|	|U|
|H|	|L|

RcvP elements		
Aperture	C	V
Location	C	V
Laryngeal (Phonation/Tone)	C	V

As discussed in the preceding section, the resort to an arboreal structure enriched with gestural nodes allows RcvP to shrink the number of elements to a binary set. Notice, however, that the gestural nodes constitute a sort of representational primitives themselves, even if of a different nature than components/elements. As a consequence, GP and RcvP display the same number of elements, the difference between the two theories consisting mostly in the presence *vs* absence, within the representational toolbox, of the gestural nodes. In other words, GP and RcvP differ in the relative balance between structure and substance: whereas GP decides to minimize structure and maximize substance (*viz* many elements on small trees), RcvP gets rid of most substance by maximally exploiting the structural dimension (*viz* a few elements on bigger trees).

Among the directions GP evolved into, variants can be found that try to reduce substance in a similar fashion[4]. One of the targets of Occam's razor is the |A| element, which is repeatedly shown to behave differently from other vocalic/resonance elements such as |I| and |U|. For instance, |A| is argued to be more syllabic than |I| and |U|, thereby showing a preference for occupying the head position of nuclei while avoiding the nuclear dependent position[5]. Furthermore, |A| is shown to interact with nasalization and length[6].

As discussed above, the price to pay for substance reduction is structure enrichment. As a consequence, |A| is replaced by structure. For instance, in order to formalize the preference for |A| syllabicity, Rennison (1998) proposes to associate the phonetic counterpart previously related to |A| (i.e. a centrally converging F1–F2 acoustic pattern) with the presence of a nuclear position lacking any elemental specification (whose unmarked status is thus representationally encoded; see Cavirani and van Oostendorp *in press* for a similar proposal).

An even more extreme development of GP towards substance reduction is represented by what came to be known as GP 2.0 (Pöchtrager 2006; Živanović and Pöchtrager 2010; Kaye and Pöchtrager 2013; Schwartz 2010), which eliminates |H|, |ʔ| and |A| by resorting to structures and mechanisms inspired by syntactic analogues, such as control[7], m-command[8] and head-adjunction. Let's focus now on |A| (referring the reader to Kaye and Pöchtrager 2013 for |H| and |ʔ|).

As just mentioned, |A| is argued to display a special interaction with length: "more specifically, |A| seemed to make bigger structure possible" (Pöchtrager 2015: 261). As a matter of fact, what is traditionally referred to as |A| is formalized as pure structure, where the extra structure is guaranteed via head adjunction: "in the case of head adjunction, the head xN projects to another level but remains the same type, i.e. an xN" (Pöchtrager 2015: 261). This is shown in (20), where the arrow between xN and its sister (in [a]) represents control:

---

4 While discussing the variants which resort to more elaborate arboreal structure to get rid of elements, Backley (2012: 75) warns that the standard theory "manages to strike a useful balance between the two, providing a restrictive model of phonological knowledge in which elements are abstract enough to function as cognitive units of linguistic structure yet concrete enough to be realized phonetically without the need for explicit rules of phonetic interpretation".

5 As discussed in Backley (2012), this could be the reason why, for instance, diphthongs such as [ai] and [au] are typologically less marked than [ia] and [ua], which is in turn possibly related to the fact that only the latter diphthongs are reinterpreted as glide-vowel sequences. This shows that, whereas |I| and |U| may be (re)interpreted as belonging to the onset preceding the |A| nucleus, the same does not hold for |A|, which keep on projecting to its nuclear node.

6 For instance, in French, where only |A| nuclei can be lengthened and nasalized (Ploch 1995).

7 In Pöchtrager (2006:77) control is described as, "[an] unannotated x in a non-maximal onset projection must be controlled by its xO [*viz* the onset head]." In GP2.0 control is generalised to structures occurring in nuclear projections. Its general effect is that of making the controlled point inaccessible.

8 Živanović & Pöchtrager (2010) define m-command as a sort of licensing necessary for phonetic interpretation, whereby terminals, i.e. elements or empty structural position, can be interpreted only if m-commanded. In the case of an empty structural position, m-command has the same effect of spreading, the commanded receiving the same phonetic interpretation of the commander.

(20) GP 2.0 vocalic elements

```
[ɪ]     [ʊ]     [ɨ]      [ə]        [a]
xN      xN      xN       xN         xN
 |       |               / \        / \
|I|     |U|           xN   x     xN➔x

   no head adjunction        head adjunction
```

As shown in (20), both [ə] and [a] are represented as pure structure. The only difference is the presence in the latter of control, which is thus deemed the responsible for the [a]-interpretation of such an empty structure, otherwise sounding [ə]. Note that “the control relationship also expresses that within [a] both positions are used up, while in [ə] there is one position (the non-head) available. In some sense, [ə] takes up less room than [a]. This neatly capture Lowenstamm’s (1996) observation that [ə] is the shorter version of [a][9]” (Pöchtrager 2015: 261). Furthermore, the lack of control in [ə] is considered to be the reason why it can be coloured by adjacent melody: the absence of control leaves “one position […] available”, which can thus host elements spreading from adjacent structures (e.g. in the analysis of Putonghua proposed in Živanovič and Pöchtrager 2010).

3.4 Comparison

As shown in the preceding sections, a research line can be identified within the DP-inspired tradition which aims at reducing substance by exploiting structure. This is particularly evident in the case of RcvP, which attempts at minimizing substance (only |C| and |V| are left) by enriching the structural dimension, whereas GP minimizes structure and maximizes substance (|ʔ|, |H|, |L|, |A|, |I| and |U|). Even within the latter, though, variants have been proposed that prefer to pay a little structural price to get rid of elements (which reduce to |I|, |U| and |L|).

Focusing on the vocalic half of the phonological world, we try to go even further by eliminating all the substantial content. More precisely, following the path initiated by DP and RcvP, we exploit the possibility for a given primitive to occur in different structural positions, the difference laying in the fact that, instead of components, we replace (“Aristotelian”) substance with pure structure, namely with treelets which are recursively nested under other treelets.

This move echoes the attempts we mentioned above to reduce e.g. |a| to |V| (Anderson and Ewan 1987: 215) or |A| to empty nuclear positions (Rennison 1998). In a similar fashion, we propose a representational account of markedness whereby, differently from Rennison (1998), [ə] is represented as single-branched treelet hosting an empty node.

Furthermore, assuming that, as proposed e.g. by Lowenstamm (1996), Pöchtrager (2006) and Živanovič and Pöchtrager (2010), [ə] is the shorter version of [a] and that “|A| seemed to make bigger structure possible” (Pöchtrager 2015: 261), we represent [a] as a binary treelet containing two nodes that, in turn, host an empty node each. In prose, this means that [a] is tantamount to two schwas. This allow us to get rid of the control mechanism introduced by Pöchtrager (2006; 2015) and Živanovič and Pöchtrager (2010) to account for the difference between two sounds—[a] and [ə]—that are otherwise represented in an identical fashion. Note that we keep something similar to head-adjunction, even if, as a matter of fact, we do not need to make any head-dependent distinction.

Together with control, we can also get rid of the c-command solution proposed by Pöchtrager (2015) to solve the problems raised by the Complexity Condition[10] (Harris 1990). In a nutshell, the concerns of Harris (1990) and Pöchtrager (2015) relate to the preference for complex elemental structures to occupy the head position of diphthongs, thus for diphthongs’ heads to contain |A|.

According to Pöchtrager (2015), the “problems [of Harris (1990) account] stem from a failure to take into account the individual nature of elements, their individual character”. As a consequence, Pöchtrager (2015) proposes the structures in (20) as well as c-control, a mechanism evidently (though ‘unfaithfully’[11]) borrowed from syntactic theory. In

9 Note, also, that control is somehow analogous to standard GP headedness, whereby controllers head controlees. See Cavirani & van Oostendorp (in press) for a slightly different proposal on the structural relatedness of [a], [ə] and empty nuclei.

10 “a. Let α and β be segments occupying the positions A and B respectively. Then, if A governs B, β must not be more complex than α;

b. The complexity value of a segment is simply calculated by determining the number of elements of which it is composed” (Harris 1990: 274).

11 As recognized by Pöchtrager (2015: 270) himself, “syntactic binding is about co-reference, while phonological bounding [is about] distributional restrictions on melody”. There seem to be other problems with binding and c-command as well, as c-command (alone) is not enough, for it needs an extra mechanism to constrain its application domain (for this reason, c-command has been ‘expanded’ into the

the present paper, rather than introducing c-command, we derive the same effect from the structural properties of ‘elements’, namely from their “individual nature”: complex structures preferably contain [a] because its representation consists of two nodes that can be further expanded by adjoining additional vocalic structures (with the limitations discussed in section 1.2 above).

Similarly, the representations we propose for |I| and |U|, whereby only the former present expandable nodes (see (6)), might account for their asymmetrical behavior. This asymmetry is also discussed in Pöchtrager (2015: 258), who claims that “the English vowel system never allows combinations of |I| and |U| within some phonological expression. [This] is true for monophthongs [and] diphthongs”[12]. As in the case of |A|, we encode this asymmetry in the representations we propose for |I| and |U|, rather than resorting to c-command, whereby “I can bind U, but U must not bind I”, where “α binds β iff α c-commands β” (Pöchtrager 2015: 263).

With respect to standard element theory (Backley 2011), a crucial difference concerns phonetic interpretation. As we mentioned above, the standard theory provides “a restrictive model of phonological knowledge in which elements are abstract enough to function as cognitive units of linguistic structure yet concrete enough to be realized phonetically without the need for explicit rules of phonetic interpretation”. Note that the more elements/substance we replace by structure, the more complex the phonetic interpretation procedure. Assuming strict modularity, though, whereby phonology and phonetics are two different realms and the former is translated into the latter in a lexical access fashion (Scheer 2014), this problem is perhaps not that dangerous. If anything, there would be an issue of learnability, which is arguably much simpler assuming that primes are at the same time cognitive *and* concrete units, as per the element standard theory.

4 Conclusion: substance as structure

From one fairly simple definition in (1), we can derive a theory that seems to mimic the results of known Element Theories. However, this theory has two properties that are seemingly disadvantageous. First, it does not seem to add a lot of new empirical coverage to what we know from Element Theory, also because the latter is fairly successful. Secondly, although the theory we present is extremely simple, it leads to representations of even simple vowels like /ɔ/ that are rather difficult to read (such as those in (17)).

In other words, our graph theoretic definition gives a notational variant of existing theories in an obtuse notation. So why would it be helpful to consider this? We believe that in most practical analyses it will be more useful to write vowels in terms of familiar |I|, |A|, |U| representations, but the set theoretic definition gives us insight into the internal structure of these elements: why there are three of them, why only one of them can fully embed, whereas the other two are heavily deficient in this respect and, possibly, why we have at most 4 degrees of vowel height. We thus get a deeper insight into the reason why elements function a certain way that would not be available if we treat them as completely primitive, atomic elements. At the same time, for studying e.g. the vowel set of a particular language, we may not always need to know why elements function in some way. This is of course familiar from most kinds of (linguistic) analysis. For instance, for the analysis of stress, we typically do not need to give the full internal structure of all vowels involved: we use the ‘higher-order’ representation of IPA symbols with the understanding that these stand for combinations of elements. On the other hand, sometimes certain features may be relevant for the assignment of stress (like height features, or tone).

Our theory is similarly a theory of vocalic elements; it aims to explain the properties of these elements, but in the everyday business of phonological analysis, it may not be necessary to refer to them all the time.

Notice that, since our system has definite properties, it is not compatible with all the possible interpretations of element theory. For instance, it is impossible to introduce a notion like ‘headedness’ into the system without making crucial changes to it. The kind of asymmetry that headedness applies can only come about by an extra theoretical device that is not available in the current theory.

c++command in Živanovič & Pöchtrager 2010). However, “the locality of binding cannot be captured by this alone, but also by no other command relationship I can envision. How this boundedness is formally captured is another question” (Pöchtrager 2015: fn 11).

12 Other asymmetries listed by Pöchtrager (2015), which can be accounted for by our representational proposal, are the fact that Turkish has a) two /e/’s but only one /o/ and b) a |U|-harmony that is more restricted than |I|-harmony.

References

Adelson-Velsky, Georgy and Evgenii Landis. 1962. An algorithm for the organization of information. *Proceedings of the USSR Academy of Sciences* (in Russian) 146, 263–266. [English translation by Myron J. Ricci in Soviet Math. Doklady 3,1259–1263].

Anderson, John and Colin J. Ewen. 1987. *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, John and Charles Jones. 1974. Three theses concerning phonological representations. *Journal of Linguistics 10*, 1–26.

Backley Phillip. 2011. *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Backley Phillip. 2012. Variation in Element Theory. *Linguistic Variation* 12.1, 57–102.

Bale, Alan, Charles Reiss and David Ta-Chun Shen. Submitted. Sets, rules and natural classes: { } vs. [ ].

Botma, Bert and Marc van Oostendorp. 2012. A propos of the Dutch vowel system 21 years on, 22 years on. In Bert Botma and Roland Noske (eds.), *Phonological Explorations: Empirical, Theoretical and Diachronic Issues*. Berlin: De Gruyter.

Cavirani, Edoardo and Marc van Oostendorp. 2017. The markedness of silence. In Bridget Samuels (ed.), *Beyond Markedness in Formal Phonology*. Amsterdam: John Benjamins, 101–120.

Charette, Monik. 1991. *Conditions on phonological government*. Cambridge: Cambridge University Press.

Chomsky Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

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

Crothers, John. 1978. Typology and universals of vowel systems in phonology. In Joseph H. Greenberg, Charles A. Ferguson and Edith A. Moravcsik (eds.), *Universals of Human Language: Phonology, Vol. 2*. Stanford: Stanford University Press, 95–152.

Harris, John. 1990. Segmental complexity and phonological government. *Phonology 7*, 255–300.

Harris, John and Geoff Lindsey. 2000. Vowel patterns in mind and sound. In Noel Burton-Roberts, Philip Carr and Gerry Docherty (eds), *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press, 185–205.

Hulst, Harry van der. 1988. The geometry of vocalic features. In Harry van der Hulst and Noval Smith (eds.), *Features, Segmental Structure and Harmony Processes*. Dordrecht: Foris, 77–126

Hulst, Harry van der. 1994. An introduction to Radical CV Phonology. In Susanna Shore and Maria Vilkuna (eds.), *SKY 1994: Yearbook of the Linguistic Association of Finland*. Helsinki, 23–56.

Hulst, Harry van der. 1996. Radical CV Phonology: The segment: syllable connection. In Jacques Durand and Bernard Laks (eds.). *Current trends in phonology: Models and methods, Vol 1*. CNRS/ESRI Paris X, 333–363.

Hulst, Harry van der. 2006. Dependency phonology. In Keith Brown (ed.), *The Encyclopedia of Language and Linguistics, 2nd edition, Volume 3*. Oxford: Elsevier, 451–458.

Hulst, Harry van der. 2011. Dependency-based phonologies. In John Goldsmith, Jason Riggle and Alan Yu (eds.), *The Handbook of Phonological Theory, 2nd edition*. Malden, MA: Wiley-Blackwell, 533–570.

Hulst, Harry van der. 2015. The opponent principle in RcvP: binarity in a unary system. In Eric Raimy and Charles Cairns (eds.), *The Segment in Phonetics and Phonology*. Chichester: Wiley-Blackwell, 149–179.

Hulst, Harry van der. 2017. The integration of segmental and syllabic structure in Radical CV Phonology. Paper presented at the workshop *The Interface Within. What relations hold between prosody and melody?*, 13 March 2017, Meertens Institute, Amsterdam, The Netherlands.

Kaye, Jonathan D., Jean Lowenstamm and Jean-Roger Vergnaud. 1985. The internal structure of phonological representations: a theory of charm and government. *Phonology Yearbook* 2, 305–328.

Kaye, Jonathan D., Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7, 193–231.

Kaye Jonathan D. and Markus Pöchtrager. 2009. *GP 2.0*. Paper presented at the "Government Phonolog Round Table 2009", Piliscsaba, Hungary.

Kaye, Jonathan D. and Markus Pöchtrager. 2013. GP 2.0. *SOAS Working Papers in Linguistics* 16, 51–64.

Lowenstamm, Jean. 1996. CV as the only syllable type. In Jacques Durand and Bernard Laks (eds.), *Current Trends in Phonology: Models and Methods*. European Studies Research Institute, University of Salford Publications, 419–442.

Nasukawa Kuniya. 2015. Recursion in the lexical structure of morphemes. In Marc van Oostendorp and Henk van Riemsdijk (eds.), *Representing Structure in Phonology and Syntax*. Berlin/Boston: Mouton de Gruyter, 211–238.

Ploch, Stefan. 1995. French nasal vowels: a first approach. *SOAS Working Papers in Linguistics* 5, 91–106.

Pöchtrager, Markus. 2006. The structure of length. Ph.D. dissertation, University of Vienna.

Pöchtrager Markus. 2012. Deconstructing A. Paper presented at the "MFM Fringe Meeting on Segmental Architecture", Manchester, U.K.

Pöchtrager Markus. 2015. Binding in phonology. Marc van Oostendorp and Henk van Riemsdijk (eds.), *Representing Structure in Phonology and Syntax*. Berlin/Boston: Mouton de Gruyter, 255–275.

Reiss Charles. 2017. Substance Free Phonology. In S. J. Hannahs and Anna R. K. Bosch (eds), *Handbook of Phonological Theory*. London: Routledge, 425–452.

Rennison, John R. 1998. Contour segments without subsegmental structures. In Eugeniusz Cyran (ed.), *Structure and Interpretation: Studies in Phonology*. Lublin: Folium, 227–245.

Schane, Sanford A. 1984. The fundamentals of particle phonology. *Phonology Yearbook* 1, 129–155.

Scheer, Tobias. 2004. *A lateral theory of phonology: What is CVCV, and Why Should It Be?* Berlin/New York: Mouton de Gruyter.

Scheer, Tobias. 2012. *Melody-free syntax and two Phonologies*. Paper presented at the Annual Conference of the Réseau Français de Phonologie (RFP), Paris, 25-27 June 2012.
Scheer, Tobias. 2014. Spell-out, post-phonological. In Eugeniusz Cyran and Jolanta Szpyra-Kozlowska (eds.), *Crossing Phonetics-Phonology Lines*. Newcastle upon Tyne: Cambridge Scholars Publishing, 255–275.
Schwartz, Geoff (2010). Auditory representations and the structures of GP 2.0. *Acta Linguistica Hungarica* 57.4, 381–397.