Dittongo mobile and other diphthongs: no allomorphy, just phonology*

Edoardo Cavirani

KU Leuven/CRISSP

Abstract. Some Italian forms display an alternation between front and back diphthongs ([jɛ]/[wɔ]) in open, stressed syllables, and mid front and mid back vowels ([e]/[o]) in other contexts. In previous literature, a couple of fully phonological accounts have been proposed. Due to their overgeneration issues, though, these analyses were abandoned, and nowadays most literature agrees in conceiving of this alternation as a case of phonologically-conditioned allomorphy. This paper challenges this view and proposes a fully phonological analysis that capitalizes on the refinement of phonological representations. This is obtained by combining a version of strict CV (Scheer 2022) with Turbidity Theory (Goldrick 2001). This systems allows for the collapsing of the differences of the relevant allomorphs in the representation of the underlying form they derive from. Crucially, the latter contains a 'turbid' glide that surfaces only in stressed, open syllables, where the lengthening of the nucleus crucially provides room for the glide to surface. As shown, this system prevents the overgeneration affecting previous proposals, and provides further support to the substance-free take on phonology.

1 Introduction

Some Italian forms display an alternation between simple vowels and diphthongs. More specifically, mid front and mid back vowels (e/o) alternate with front and back diphthongs (jɛ/wɔ) depending on their phonological context: diphthongs can only be found in open stressed nuclei, whereas in all the other contexts the corresponding single vowel appears. This is illustrated in Table 1 with the PRS.IND forms of *sedere* 'to sit' (left-

* This work is funded by FWO Marie Skłodowska-Curie Actions - Seal of Excellence postdoctoral research grant nr.12Z7520N. Previous versions of this work have been discussed with Michal Starke, Maria Cortiula and other members of the *nanolab*, and with Nicola Lampitelli, whom I thank for many valuable comments.

hand side) and *morire* 'to die' (right-hand side). As can be seen, stressed diphthongs (PRS.IND.1/2/3SG/3PL) alternate with unstressed monophthongs (PRS.IND.1/2PL).

	SG	PL	SG	PL
1	ˈsjɛdo	seˈdjaːmo	ˈmwɔjo	moˈrjaːmo
2	ˈsjɛdi	seˈdeːte	ˈmwɔri	moˈriːte
3	ˈsjɛde	ˈsjɛdono	ˈmwɔre	ˈmwɔjono

Table 1: *Dittongo mobile* - the role of stress

The relevance of syllable structure is shown by the examples in Table 2. Differently from the forms in Table 1, the PRS.IND.1SG and PRS.IND.3PL forms of *venire* 'to come' (left-hand side) and *potere* 'to be able to' (right-hand side) show no diphthongs. Crucially, the relevant nuclei bear stress, but they occur in closed syllables, which prevents lengthening.

	SG	PL	SG	PL
1	ˈvɛŋgo	veˈnjaːmo	ˈpɔsːo	posˈsjaːmo
2	ˈvjɛni	veˈniːte	ˈpwɔi	poˈteːte
3	ˈvjɛne	ˈvɛŋgono	ˈpwɔ	ˈpɔsːono

Table 2: *Dittongo mobile* - the role of syllable structure

These alternations can mainly be found in verbal paradigms, where forms are clearly related to each other. In nominal derivational paradigms, similar alternations can also be found, e.g. [ˈwɔmo] 'man' ∼ [ˈomiːno] 'little man', but in most cases it is not quite clear whether the forms involved share one and the same root, e.g. [ˈdjɛtʃi] 'ten' ∼ [deˈtʃiːna 'ten or so'].[1]

In the literature, these alternating diphthongs are known as *dittongo mobile* 'mobile diphthong', and have attracted considerable attention from Italian language historians (see van der Veer 2006 for an overview). However, in the recent phonological literature, *dittongo mobile* has received little attention, as only a couple of analyses have been proposed. These analyses either conceive of *dittongo mobile* as a case synchronic phonological process (Saltarelli 1970; Sluyters 1992), or as a case of phonologically-conditioned allomorphy (van der Veer 2006; Booij & van der Veer 2015). The main difference between these two kinds of analyses is that the fully phonological ones assume one underlying form and a set of rules deriving the observed surface realizations,

1. When the two forms most likely share the same root, no alternation is observed, e.g. [ˈdjɛtʃi] 'ten' ∼ [djeˈtʃiːno] '10 cent coin', [ˈljɛto] 'happy.M.SG' ∼ [ljetaˈmente] 'very happy.ADV', [ˈbwɔno] 'good.M.SG' ∼ [bwoˈnisːimo] 'very good.M.SG'.

whereas the allomorphy-based analyses argue for two different underlying forms and a set of phonological constraints that decide which form will eventually surface. Let us look at these analyses in more detail.

As just mentioned, the two fully phonological analyses maintain that the two forms involved in *dittongo mobile* derive from one and the same underlying form, but they differ with respect to which form is argued to be the underlying one. For Saltarelli (1970), the underlying forms display /iɛː/ and /uɔː/. A rule would then change high vowels into glides when followed by another vowel (/iɛː/ → /jɛː/, /uɔː/ → /wɔː/), and another rule would apply to the output of the previous rule that deletes the glides (/jɛː/ → /ɛː/, /wɔː/ → /ɔː/). Another couple of rules would then raise and shorten the vowels (/ɛː/ → /e/, /ɔː/ → /o/). Besides the fact that this proposal resorts to technical devices that in current generative phonology approaches are no longer in use (e.g. rule ordering and flat representations), it crucially fails in accounting for why the proposed derivation does not apply to all forms displaying /jɛ/ or /wɔ/. For instance, this system wrongly predicts [pjɛgo] 'I fold' ∼ *[peˈgjaːmo] 'we fold' (cf. [pjeˈgjaːmo] 'we fold') and [ˈswɔno] 'I play (music)' ∼ *[soˈnjaːmo] 'we play (music)' (cf. [swoˈnjaːmo] 'we play (music)').[2]

Sluyters (1992) proposes instead that the underlying form of *dittongo mobile* is the one with a monophthong. A diphtongization rule would then apply to these forms deriving the observed diphthongs. The trigger of this process would be a constraint requiring feet to be binary in open stressed syllables, which is the same context triggering vowel lengthening. As in the case of Saltarelli (1970), this system also overgenerates and wrongly predicts unattested patterns, such as [beˈvjaːmo] 'we drink' ∼ *[ˈbjɛːvo] 'I drink' (cf. [ˈbeːvo] 'I drink') and [voˈtjaːmo] 'we vote' ∼ *[ˈvwɔːto] 'I vote' (cf. [ˈvoːto] 'I vote').

Thus, the main issue with the fully phonological analyses proposed so far is that they overgenerate and predict wrong forms, namely, they predict that all the forms displaying [ˈjɛ]/[ˈwɔ] alternate with forms with [e]/[o] (Saltarelli 1970), and that all the forms displaying [e]/[o] alternate with forms with [ˈjɛ]/[ˈwɔ] (Sluyters 1992). This is not what we observe, as there are plenty of forms that keep the monophthong or the diphthong throughout the whole inflectional and/or derivational paradigms. The fact that *dittongo mobile* alternations can only be found in a subset of forms has been thus taken as an argument favouring an allomorphy-based account: as synchronic phonology doesn't seem to be able to derive the lexical restrictions on *dittongo mobile*, the alternating

2. While Saltarelli (1970) maintains that the vocalic part of the diphthong is long, recent acoustic analyses by van der Veer (2006) show that the whole duration of stressed diphthongs is comparable to that of stressed monophthongs, thus [ˈjVː]=[ˈVː]. In what follows, we will stick to and build on this experimental evidence.

forms must be stored in the Lexicon as such.

This is the take of van der Veer (2006) and Booij & van der Veer (2015), who claim that "[dittongo mobile] is not a phenomenon triggered by active phonological processes but an instance of mixed phonological and morphological allomorphy". They "posit multiple stems for each morpheme, so, for instance, the verb *sedere* would have two underlying allomorphs: /sɛd/ and /sjɛd/ [, and] the selection of either /sɛd/ and /sjɛd/ is predicted by the interaction of universal faithfulness and markedness constraints" (van der Veer 2006: 104). The hypothesis that the selection of the observed allomorph depends on the interaction of phonological constraints obviously grants an important role to phonology. However, to the extent that explaining means predicting (Egré 2015), and that predicting is only possible within a constrained formal system, listing allomorphs in the Lexicon is tantamount to introducing exceptions into a system, which clearly translates into a decrease of the explanatory power of a theory. Allomorphy should thus be considered a last resort strategy, which is called upon only if it is shown that the system fails in generating the observed surface forms from a single input. Crucially, whereas reducing two different surface forms to a single underlying form is sometimes utterly impossible, there might be cases in which refining the computational (e.g. Zimmermann 2017) or the representational (e.g. Scheer 2017) component of phonology would allow us to dispense with allomorphy. This paper makes a similar case, inasmuch as it offers a fully phonological analysis of *dittongo mobile* that capitalizes on the refinement of the phonological representation of the relevant underlying forms. This is obtained by combining a version of strict CV (Lowenstamm 1996; Scheer 2004; 2022) with Turbidity Theory (Goldrick 2001; Cavirani & van Oostendorp 2017; 2020; Cavirani 2022). This systems allows for the collapsing of the representational differences of the *dittongo mobile* allomorphs in the representation of one and the same underlying form. In a nutshell, rather than having two allomorphs, one with a glide and another without glide, I postulate one single underlying form that contains the glide. Crucially, the glide surfaces only in a specific context, namely when the nucleus containing it occurs in a stressed, open syllable. In such a context, the lengthening of the nucleus provides room for the glide to surface. Note that, as mentioned above (fn. 2), the duration of stressed diphthongs equals that of stressed monophthongs. This means that the vocalic part of the diphthong is not lengthened. This contrasts with what is observed in non-alternating monopthongs. In this case, I argue that the underlying representation contains no glide, thus the extra room provided by stress can be fully 'used' by the vowel, which lengthens. A still different scenario is that of non-alternating diphthongs, which can be either short or long. The difference between non-alternating and alternating diphthong lies in the status of the glide. More precisely, while in non-alternating diphthongs the glide is associated with the nucleus in the same way the vocalic part of the diphthong is, in alternating diphthongs the association of the glide is 'weaker'.

In what follows, I will introduce the theoretical assumptions and the formal toolkit necessary for the analysis (section 2), and I will work out the details of the latter (section 3). This will clarify what is meant by 'weaker' association. In the conclusion (section 4), I will also briefly hint at the possibility of extending the proposed representational format to virtual geminates.

2 Theoretical toolkit

2.1 Strict CV

The first set of assumptions I adopt for the analysis of *dittongo mobile* come from a strict CV (Lowenstamm 1996; Scheer 2004). According to this theoretical approach, all phonological strings consist of CV sequences. Thus, what superficially looks like a geminate (Fig. 1a), or a consonant cluster (Fig. 1b), will phonologically consist of a CVCV sequence whose left-hand V has no melodic content. The same goes for long vowels (Fig. 1c) and hiatuses (Fig. 1d), which are conceived of a CVCV sequence with an internal empty C.

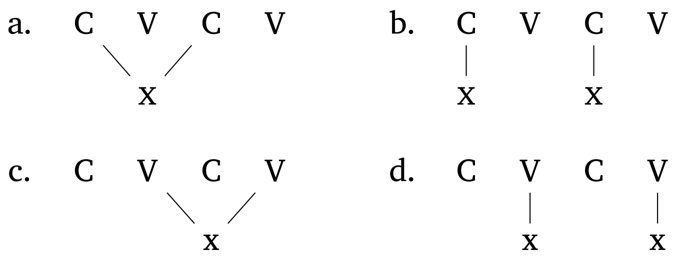

Figure 1: Strict CV structures

As strict CV doesn't recognize any autonomous/formal status to the syllable (nor to the rhyme), patterns that can be explained with reference to syllabic structure (e.g. open syllable lengthening, closed syllable shortness, extrasyllabicity, coda effects, positional strength, vowel ∼ ∅ alternations, etc.) must be given an alternative explanation. Given the flat structure of strict CV representations, this explanation must be found in the lateral relations established between the CV nodes. These relations are assumed to be a function of the forces discharged by full V nodes, which always apply leftward and come in two flavours: government and licensing. The former is conceived of as a force that aims at silencing its target. If the latter is empty, it can stay silent (e.g. no epenthesis), whereas if it contains some melodic structure, such structure is reduced (e.g. lenition). Licensing works the other way around, as it supports the pronunciation of the melodic content of its target.[3] These relations are represented as in Fig. 2, where the arrow on top of the CV tier stands for government, and the one at the bottom stands for licensing.

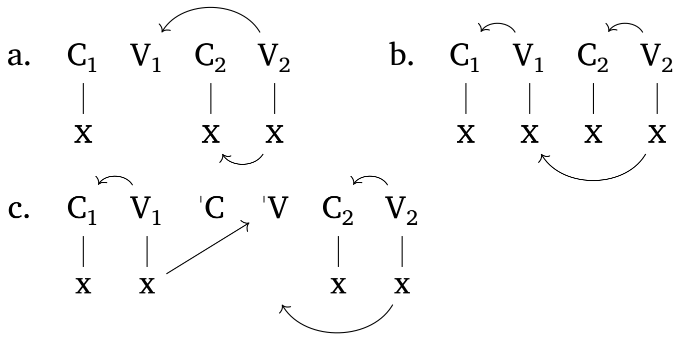

Figure 2: Strict CV lateral forces

In a fairly standard version of strict CV (Scheer 2004) the distribution of the two lateral forces depends on two principles: i) a position cannot at the same time be governed and licensed, and ii) government takes precedence over licensing. Another principle seems to be required by this version of strict CV that requires a full V to discharge both forces, as it is claimed that government "goes to the preceding empty nucleus, or to the preceding onset if the preceding nucleus is not empty [and] licensing, in turn, affects the position that escaped government" (Scheer & Cyran 2018: 274).[4] Let us see how the system works in Fig. 2.

In Fig. 2a, the full V_2 scans the string leftward, finds the empty V_1, and discharges its government force on it. V_1 can thus stay silent and, by virtue of being governed, it cannot itself discharge any force. Once its government force is discharged, V_2 can discharge its licensing force on C_2. As for C_1, it is neither governed nor licensed, as the following governed V cannot discharge any lateral force. In Fig. 2b, V_1 is full, so V_2 discharges its government on C_2, and its licensing on V_1 (which in turn governs C_1, as there is no preceding empty V). In Fig. 2c, we see

3. Depending on the specific version of strict CV one adopts, these two basic forces can be upgraded and/or adapted to specific contexts, or reduced to one (see Scheer & Cyran 2018 for an overview).

4. This further suggests that a full V automatically discharges its forces on the preceding VC nodes without any trigger, and that, if the preceding V and C are both full, government targets C. As far as I know, these extra assumptions are not discussed in the literature.

an extra CV sequence: ˈCˈV. According to Larsen (1998), this extra CV is introduced by stress, and provides the space for the lengthening of the stressed V, represented as the spreading (rightward arrow) of the melodic content of V_1 to ˈV. Given that ˈV is associated with the melodic content of V_1, it is not empty. As a consequence, V_2 discharges its government force on C_2 and its licensing force on ˈV. According to Larsen (1998), this is the reason why stressed vowel lengthening can only occur in open syllables. In other words, for lengthening to happen, the (content of the) extra V must be licensed.. As we will see, this hypothesis plays a crucial role in the analysis.[5]

2.2 Turbidity Theory

Another crucial formal tool for my analysis is Turbidity Theory (TT; Goldrick 2001; van Oostendorp 2008; de Castro-Arrazola et al. 2015; Torres-Tamarit 2015; Cavirani & van Oostendorp 2017; 2019; Cavirani 2022). TT builds on a version of Optimality Theory (Prince & Smolensky

5. Note that for Larsen (1998)'s hypothesis to hold, we need to order the spreading operation responsible for the lengthening of the stressed V before the discharge of the lateral forces of the following V. This might not be a trivial issue for several reasons. Let us consider an example from Italian, e.g. [ˈkaːsa] 'house'. Under relatively standard assumptions, the derivation of this form starts from the selection of the root /kas-/, which is then merged with the GENDER and NUMBER heads, realized by /-a/ if values are FEMININE and SINGULAR (I gloss over the discussion concerning the internal morphosyntactic structure of /-a/$_{\text{F.SG}}$). In strict CV, the root /kas-/ is represented as a CVCV string with a final empty V, and /-a/$_{\text{F.SG}}$ as a floating segment (Passino 2009; Lampitelli 2010). Given that the Italian stress assignment algorithm places the stress on penultimate vowels, and that this algorithm can only see full V nodes, the stress gets assigned to the first V node of /kasØ/ only after the floating /a/$_{\text{F.SG}}$ associates to the final empty V node of the root. Thus, we would have /kasØ/+/a/$_{\text{F.SG}}$ → /kasa/$_{\text{F.SG}}$ → /ˈkasa/$_{\text{F.SG}}$. According to Larsen (1998), when the stressed V is in an open syllable, i.e. when it is followed by a full V, the extra *empty* CV responsible for the lengthening of the stressed V is introduced. We would thus have /ˈkasa/$_{\text{F.SG}}$ → /ˈkaØ$_{\text{C}}$Ø$_{\text{V}}$sa/$_{\text{F.SG}}$. Note though, that at this point the extra V is i) empty and ii) followed by a full V. Thus, the latter could in principle govern the former, forcing it to stay silent, yielding *[ˈkasa]. Conversely, we would expect 'closed syllable lengthening', as an empty V node cannot govern a preceding empty V, i.e. /kasØtØ/+/a/$_{\text{F.SG}}$ → /kasØt-a/ → /ˈkaØ$_{\text{C}}$Ø$_{\text{V}}$sØt-a/ → *[ˈkaːsta] (cf. [ˈkasta]) 'chaste$_{\text{F.SG}}$'. For the system not to generate the wrong forms, we necessarily have to maintain that the melodic content of the stressed V fills in the extra V provided by stress before the final full V discharges its governing force (remember that, by assumption, government takes precedence over licensing). Only in this case, the final V's government force can spare the preceding V and target C. However, stress can be assigned to the relevant V only if the following V is not empty. It would thus look like we have to assume that the final V is full (to derive the correct stress pattern), but it must wait to discharge its lateral forces. More work is needed to solve this issue. For the time being, I assume the standard, albeit possibly perfectible, hypothesis of Larsen (1998).

1993) that considers the input-output relation as a containment relation, i.e. it assumes that the input is always contained in the output. This means for instance that deletion is conceived of as phonetic underparsing of a melodic prime, rather than as the complete removal of the latter from the phonological structure. This formalization is granted by the splitting of the symmetric autosegmental relation in two components: a projection relation expressing the lexical affiliation of a melodic prime to a given prosodic node, and a pronunciation relation expressing the phonetic interpretation of a melodic prime in a specific prosodic node. These two relations are different from each other also with respect to their status and availability to phonological computation. Projection relations are part of the underlying representation, and cannot be altered because of Containment and Consistency of Exponence (van Oostendorp 2008), whereas the pronunciation relations can be modified, namely they can be added, removed, or shifted depending on the requirements of the phonological system, e.g. under the pressure of structural constraints holding on surface representations.[6] Conventionally, projection and pronunciation relations are depicted as arrows: an arrow pointing from the prosodic node to the melodic prime represents the projection relation, whereas an arrow pointing from the melodic prime to the prosodic node represents the pronunciation relation. This is illustrated in Fig. 3. Fig. 3a represents a floating segment, namely a segment that is associated to no prosodic node; Fig. 3b represents an empty prosodic node, which has no melodic content, hence no relation either; Fig. 3c represents a silent non-empty prosodic node, i.e. a node that has some melodic content that is not pronounced (viz it only has the projection relation); Fig. 3d represents a full prosodic node, which has some melodic content that is associated with its prosodic node via both a projection and a pronunciation relation (for the sake of clarity, in what follows, I have silent segments coloured in light gray).

a.	b. C/V	c. C/V	d. C/V
		↓	↕
x		x	x

Figure 3: Turbidity Theory representations

Phonetically, the melodic content of Fig. 3d is faithfully interpreted, as dictated by its pronunciation relation. As for the other structures, their phonetic interpretation depends on the context in which they occur. If

6. See Polgárdi (1999); Rowicka (1999); Harris & Gussmann (2002); Cavirani (2015); Faust & Torres-Tamarit (2017), and Carvalho (2019) for proposals translating strict CV mechanisms into constraint interaction.

they are followed by a full V, they can be governed, thus stay silent, otherwise they get pronounced. Unless some other process applies (e.g. spreading), empty structures such as that in Fig. 3b are assigned some default epenthetic melodic content, while silent non-empty structures (Fig. 3c) have their melodic content faithfully interpreted. As for floating segments, they get pronounced if associated to some prosodic node. In all these cases, the phonetic interpretation is formalized as the introduction of a pronunciation relation.

3 Analysis

3.1 Representations

The formal tools described in section 2.1 and 2.2 allow us to propose an analysis of *dittongo mobile* where the alternation of the attested surface forms derive from one and the same underlying form via synchronic phonological computation (section 3.2). This approach also allows us to formally distinguish between alternating diphthongs, non-alternating diphthongs, and ‘fake’ diphthongs. The representations of these objects are given in Fig. 4. In Fig. 4c we have ‘fake’ diphthongs, namely sequences of segments that should rather be conceived of as CV sequences, as opposed to ‘true’ diphthongs, which are represented either as complex segments, i.e. as segments containing two different components (Fig. 4a), or as two segments associated with the same V node (Fig. 4b). In Fig. 4, T stands for any consonant, A for any vowel (see Marotta 1988 for the lack of restrictions on the combination of glides and vowels in non-alternating diphthongs), x for segment/root node and • for the subsegmental component.

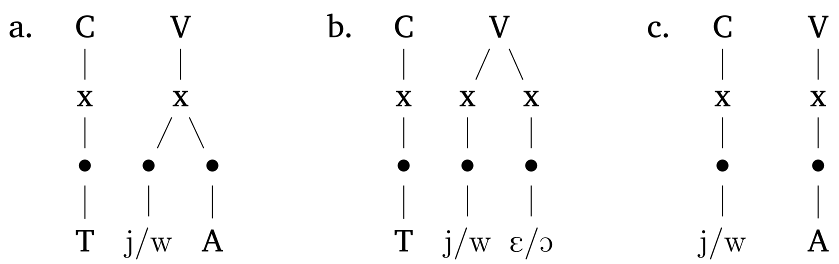

Figure 4: Diphthongs

The representations in Fig. 4 look like a departure from standard strict CV (Scheer 2004), as they feature the extra x and • nodes, whereas standard approaches only have C, V, and melodic primes. However, the need for extra representational tiers between the one hosting the prosodic nodes and the one hosting the melodic primes has been repeatedly recognized, especially within the literature dealing with complex segments, even though rarely fully formalized (see Scheer 2012 for some discussion of different approaches to complex segments). For instance, the need for root nodes in strict CV has been recently argued for by Scheer (2022), who explicitly points out that "timing units (x-slots, Cs and Vs [...], moras etc.) are not the same thing as root nodes [for the] latter have no timing properties but rather define segments" Scheer (2022: fn. 4). Earlier on, Kaye (1981; 1985) concedes that segments can have an internal complex structure, as he represents 'light diphthongs' as segments (x) that are associated to a nuclear position (N) and contain two distinct 'melodic expressions' (thus similarly to Fig. 4a). Subsegmental 'melodic expressions', or 'components', are referred to also by Rennison (1998). When describing complex segments, he claims that they consist of two components, a 'lazy' one that is realized last and is either floating, or acquired from the context, and another, stable component that is specified as such in the Lexicon. Complex segments are dealt with also by Lowenstamm (2003), who maintains that some *muta cum liquida* clusters might be represented as a "bisegmental complexes", i.e. 'C^x', "where x [...] stands for secondary articulation". The need for refining the internal architecture of segments has been recently defended also within theoretical approaches that mostly focus on computation, e.g. the OT-implemented Q-Theory (Garvin, Lapierre & Inkelas 2018; Shih & Inkelas 2019), where each segment is argued to contain three linearly-ordered subsegments (q), roughly corresponding to the closure, target, and release components of segments. Crucially, in all these cases, the fact that phonological computation is sensitive to subsegmental components (viz 'melodic expressions', 'secondary articulations', 'q', etc.) suggests that we should be able to refer to and manipulate each of them as a unit (as already pointed out by e.g. Anderson 1976). Our phonological representations should thus be endowed with a way to encode the subsegmental components, the (fact that they belong to specific) segments, and the (anchoring of segments to the skeleton, represented in terms of) CV nodes. In Fig 4, these are represented by •, x and C/V, respectively.[7]

A few empirical arguments can be mentioned that support these representations for Italian diphthongs. For instance, the representation of non-alternating diphthongs as one segment that is associated to a V node

7. The presence of • in segments without subcomponents might look superfluous. I keep it in the representations for consistency, i.e. in order to have a formally homogeneous representational format. A possibly even more consistent format would be one in which all segments have a fixed number of subsegmental components, as argued for in Q-Theory. I leave this for future research.

and contains two components (Fig. 4a)[8] can account for the fact that, despite their melodic profile, they behave like any other simple vowel, which can be pronounced as short or long depending on its phonological context. Indeed, as shown by van der Veer (2006), diphthongs are as short as monophthongs when unstressed, and as long as lengthened monophthongs when in stressed, open syllable. If the glide were a distinct vocalic segment or part of the preceding onset, we would predict that the following vocoid behaves like any other vowel. Namely, we would expect it to lengthen in stressed, open syllables. As a consequence, in such context, we would expect that its duration - rather than that of the whole diphthong - matches that of monophthongs, contrarily to what we observe.

The representation in Fig. 4a is also supported by a few of distributional arguments. Word-internally, diphthongs can be found after complex consonant clusters, as in [setːenˈtrjone] 'north.M/F.SG', [avːjaˈmento] 'start.M.SG', [cirkwiˈtaːle] 'of a circuit.M/F.SG', [iɱvjabiliˈta] 'sendability.F.SG', [consweˈtudine] 'habit.F.SG', [afːetˈtwozo] 'tender.M.SG', etc. If the first part of the diphthong were conceived of as belonging to a C node, we would generate ill-formed structures, as they would violate the two-consonant constraint holding on Italian onsets (in strict CV, the ill-formedness of these structures would derive from there being too many empty V nodes in a row, e.g. */setØtenØtØrØjØone/ [setːenˈtrjone]). Further distributional arguments (from Canalis 2018) come from i) intrinsically long consonants, which surface as long in intervocalic position and, crucially, when occuring between a vowel and a diphthong (e.g. *azione* [atˈtsjoːne] 'action', *razziamo* [ratˈtsjaːmo] 'we plunder') and ii) /s/ voicing, which, in northern Italian varieties, applies in exactly the same contexts (N.It. [ˈriːzo] 'rice' ∼ [riˈzjera] 'rice paddy', cf. St. It. [ˈriːso] ∼ [riˈsjɛːra]).

However, there are also arguments supporting an analysis in which diphthongs are represented as CV sequences (Fig. 4c). The main one comes from SG.M.DEF/INDF article selection: while words starting with a vowel select for [l] and [un] (e.g. [l ubrjaˈkoːne] 'the boozer', [l ikonoˈklasta] 'the iconoclast', and [un ubrjaˈkoːne] 'a boozer', [un ikonoˈklasta] 'an iconoclast'), words starting with a glide behave like words starting with /sC(C)/ clusters and (intrinsic) geminates, and select for [lo] and [uno] (e.g. [lo jelˈlaːto] 'the unfortunate (man)', [lo

8. Besides the work by Kaye (1985; 1981), a representation for diphthongs along these lines has been proposed by Marotta (1988) (for /wɔ/), Sluyters (1992) (for /wɔ/ and /jɛ/), van der Veer (2006), and Canalis (2018). A different view is held by Krämer (2022) and Passino, de Carvalho & Scheer (2022). Marotta (1988) also proposes a representation similar to that in Fig. 4c for /jɛ/. The main difference between the representations proposed by most of these authors and mine is that the former resort to moras and/or rhymes.

skrit'toːre] 'the writer', [lo ɲ'ɲɔːmo] 'the gnome', [uno jel'laːto] 'an unfortunate (man)', [uno skrit'toːre] 'a writer', and [uno ɲ'ɲɔːmo] 'a dwarf'). Things seem to be less clear in the case of the back glide, as words starting with /w/ can select either articles (e.g. ['l wɔmo] 'the man', [u'n wɔmo] 'a man' *vs* [lo (w)'waːdi] 'the wadi', [uno (w)'waːdi] 'a wadi' *vs* [il wi'kɛnd] 'the weekend', [uŋ wik'ɛnd] 'a weekend'). However, apart for two lexemes (*uomo* 'man' and *uovo* 'egg'), most nouns, including borrowings, seem to select for the preconsonantal articles.[9]

Thus, it seems that, depending on its position within the word, diphthongs can have two different phonological structures: word-medially, they are represented as complex nuclei (Fig. 4a), whereas word-initially they correspond to CV sequences (Fig. 4c).[10]

As for the representation of alternating diphthongs, we need one that allows for encoding the fact that when occurring in stressed, open syllables, they are as long as non-alternating diphthongs, but if unstressed or in closed syllables, they surface without the glide. Furthermore, we want to explicitly formalize the fact that they belong to the V node, rather than to the preceding C node. TT provides us with the necessary formal tools. We should thus update the representations in Fig. 4 accordingly, which gives us those in Fig. 5.

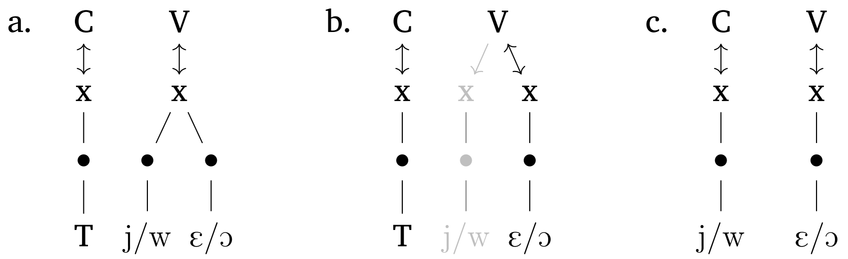

Figure 5: TT diphthongs

The representations of non-alternating (Fig. 5a) and 'fake' diphthongs (Fig. 5c) are as in Fig. 4, the only difference being that here they associate with their CV nodes via both the projection and the pronunciation relation, formalizing the fact that they are always pronounced. The

9. The literature on Italian article selection is quite vast and heterogeneous, and features both analyses that build on (phonologically-conditioned) allomorphy (Davis 1990; Marotta 1993; Tranel & Del Gobbo 2002; McCrary 2004; Russi 2006; Krämer 2009), and analyses that assume one underlying form and a set of phonological rules/constraints deriving the surface forms (Muljačić 1971; 1974; Petrosino 2018; Repetti 2020). Among the latter, see Larsen (1998); Faust, Lampitelli & Ulfsbjorninn (2018) for analyses that build on strict CV representational assumptions.

10. As a matter of fact, the fact that glide-initial nouns select for the preconsonantal article that occurs before /sC(C)/ clusters and geminates suggests that initial glides might be phonologically long.

representation of the alternating diphthong differs from the other two in two respects. Firstly, rather than as a segment with two components, it is represented as a 'bisegmental complex' (Lowenstamm 2003), namely as two segmental structures belonging to one and the same V node. In TT, 'belonging' is formally expressed by the projection relation associating a C/V node with its segmental content. Secondly, the two segments are asymmetric, in that the glide is underlyingly endowed only with the projection relation, whereas the vocalic segment has both relations. This encodes the idea that, whereas the vocalic part of the diphthong is always pronounced, the glide only surfaces if phonological computation adds the necessary pronunciation relation.[11] As we will see in the next section, alternating diphthongs are represented as 'bisegmental complexes' rather than as segments containing two distinct components because this allows us to provide for a better account of why the glide surfaces only in specific phonological contexts. In a nutshell, assuming that i) stress provides an extra CV (Larsen 1998), ii) C/V nodes can only pronounce one segment (whereas the reverse is not true, as geminates and long vowels can be understood as one segment pronounced on two C/V nodes), and iii) projected segments surface if (not governed and) there is some C/V host, the extra V provides room for rearranging the pronunciation relations in such a way that the glide is pronounced on the V node from which it is projected, and the vocalic segment on the extra V node. If we represent the glide as a subsegmental component, we wouldn't be able to capture the relation between its pronunciation and the phonological context where this happens.

3.2 Computation

As I mentioned in the previous section, if we assume that i) stress provides an extra CV (Larsen 1998), ii) C/V nodes can only pronounce one segment, and iii) projected segments surface if (not governed and) there is some C/V host, the representations in Fig. 5 allows for deriving the different behaviors of alternating and non-alternating diphthongs. Furthermore, this allows for conceiving of the alternating forms of *dittongo mobile* as deriving from one and the same underlying form via phonological computation, thus not as a case of allomorphy. This is shown in Fig. 6, where a form containing the alternating diphthong - /sjɛd/ 'sit' (Fig. 6a) - surfaces with or without the diphthong depending on the

11. This clearly reminds of the distinction Rennison (1998) makes between the 'lazy' component of complex segments, which are "either floating or acquired from the context", and the stable component, which is "specified as such in the Lexicon". In TT, 'stability' is formalized as the presence of both relations, and 'laziness' as the absence of the pronunciation relation.

phonological context. More specifically, it surfaces as [sjɛd] when the diphthong occurs in an open, stressed syllable, as in /sjɛd/ + /o/$_{\text{IND.PRS.1SG}}$ [ˈsjɛdo] 'I sit' (Fig. 6b), and it surfaces as [sed] in other contexts, as in /sjɛd/ + /ete/$_{\text{IND.PRS.2PL}}$ [seˈdeːte] 'you sit' (Fig. 6c; the additional process of raising targeting unstressed vowels is orthogonal to the present discussion, and is thus not dealt with). Let us look more closely at the two surface forms.

In Fig. 6b, we have a form resulting from the concatenation of the root - /sjɛd/ - and the phonological exponent of IND.PRS.1SG - /o/ -, which is represented as a floating segment (hence with no 'proprietary' V node and no projection relation) that associates to the root-final V_2 node via the pronunciation relation.[12] Once this happens, stress is assigned to the penultimate V_1 and the extra ˈCˈV is inserted. This configuration triggers open syllable lengthening, represented as the insertion of a pronunciation relation from the closest segmental content of V_1 to ˈV. As for the insertion of the pronunciation relation from the glide projected by V_1 to V_1 itself, I maintain that it is motivated by the constraint favouring the pronunciation of projected segments, and by the hypothesis that this constraint can be satisfied if there is a possible host for the pronunciation of the relevant segment. The constraint against the pronunciation of two segments on the same C/V node is satisfied by removing the pronunciation relation from /ɛ/ to V_1. This doesn't violate the constraint favouring the pronunciation of projected segments, as /ɛ/ is pronounced in ˈV.[13]. The resulting surface is thus [ˈsjɛdo].

Things are different if /sjɛd/ is concatenated with /ete/$_{\text{IND.PRS.2PL}}$,

12. For a fully developed formalization of Italian verbal inflectional suffixes, as well as for the interaction of *dittongo mobile* and verbs displaying a *g* infix (Tab. 2), see Cavirani, Cortiula and Starke *in prep*. There, the phonological exponent of IND.PRS.1SG is represented as a CV structure, where C is empty and V contains /o/. This allows for the surfacing of the root-final floating /g/, which can be pronounced on the empty C. With roots that have no final floating segment, the suffix's C remains empty. In the representations in Fig. 6, I represent /o/$_{\text{IND.PRS.1SG}}$ as a floating segment for expository reasons. An alternative analysis of Italian verbs' /g/ infix has been proposed by Lampitelli (2017). Although our analyses differ with respect to the source of the infix, the shape of the inflectional markers, and the derivation of the surface forms, also Lampitelli (2017) notices the interaction between the presence of the infix and the absence of the diphthong, and suggests that the alternation between diphthong and monophthong is a matter of space. However, the representation of these diphthongs, which I maintain being crucial for the analysis of the alternation and of the difference between alternating and non-alternating diphthongs, is not fully worked out. This work can thus be seen as complementary to Lampitelli (2017)'s.

13. In representation-based frameworks such as strict CV, processes can be conceived of as repairing strategies triggered by ill-formed structural configurations, which could in turn be conceived of as structural constraints holding on surface representations; see fn. 6 for relevant literature.

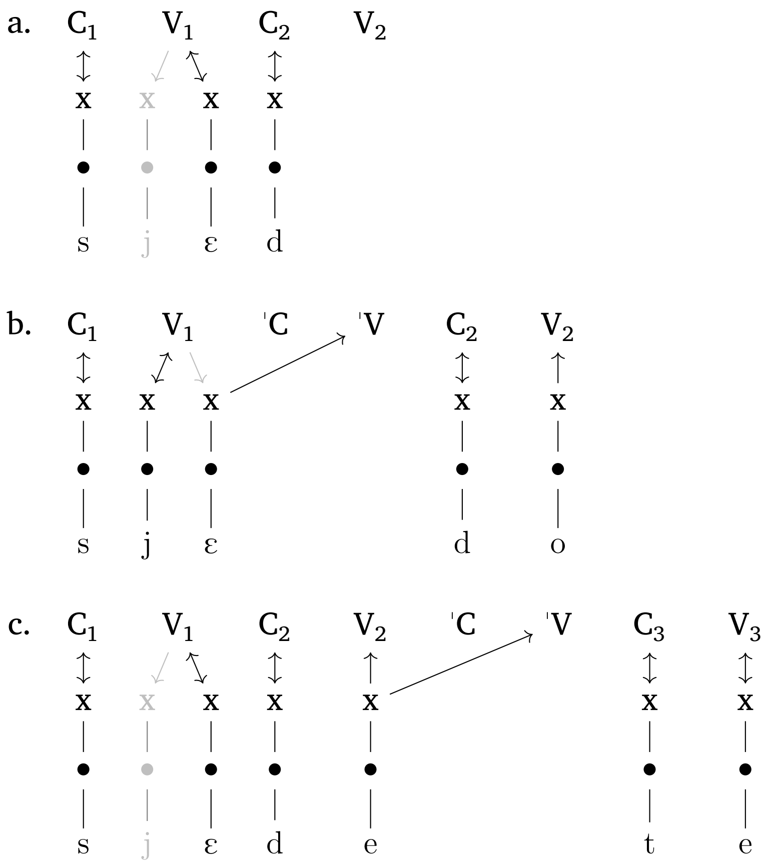

Figure 6: TT alternating diphthongs

which contains two vocalic segments (Fig. 6c).[14] Also in this case, stress is regularly assigned to the penultimate V node. Here, though, this is not the V node containing the diphthong, but V_2. The extra CV is thus inserted after V_2, triggering the lengthening of its segmental content, i.e. the insertion of a pronunciation relation from the latter to ˈV. Crucially, as no extra empty V node follows V_1, /ɛ/ cannot be pronounced anywhere else than there, thereby preventing the addition of a pronunciation relation from /j/ to V_1, which remains silent. This results in [seˈdeːte] 'you sit'.

Note that in forms with alternating diphthongs in which the latter occur in a close, stressed syllable, the glide does not surface. This is the case of verbs such as those reported in Tab. 2, repeated here in Tab. 3.

As we can see, whereas the glide does surface in IND.PRS.2/3SG, it does not in IND.PRS.1/2PL and, crucially, in IND.PRS.1SG/3PL. In the

14. The first vowel is represented as a floating segment pronounced on the root-final empty V_2 node; the quality of this floating 'thematic' vowel depends on the verb class: it is /e/ in the II class, /i/ in the III, and /a/ in the I, as in Fig. 7b.

	SG	PL	SG	PL
1	ˈvɛŋgo	veˈnjaːmo	ˈpɔsːo	posˈsjaːmo
2	ˈvjɛni	veˈniːte	ˈpwɔi	poˈteːte
3	ˈvjɛne	ˈvɛŋgono	ˈpwɔ	ˈpɔsːono

Table 3: *Dittongo mobile* - the role of syllable structure

latter, the V node hosting the diphthong is stressed, but it occurs in a closed syllable. The fact that the glide does not surface is related to the fact that stressed vowels do not lengthen in closed syllables, which, in strict CV terms, means that no extra ˈCˈV is added after the V node hosting the diphthong. As a consequence, no extra room is provided for shifting the pronunciation relation away from the latter, hence no room is provided for the surfacing of the glide.

Let us now see what happens with roots that contain a non-alternating diphthong. As discussed in the preceding section, these diphthongs differ from alternating diphthongs inasmuch as in non-alternating diphthongs the two vocoids are represented as two components of one and the same segment. This is repeated in Fig. 7a, where the underlying representation of the root /pjɛg/ 'fold' is given. In this form, /j/ and /ɛ/ belong to the segment associated with V_1. Crucially, the relation of these two components with their segment is symmetric, and the segment is associated with its V node via both a projection and a pronunciation relation. This encodes the fact that this segment and its components are always pronounced, and that the segment as a whole surfaces as short or long depending on the phonological context. This is shown in Fig. 7b and Fig. 7c. In the former, $/\text{o}/_{\text{IND.PRS.1SG}}$ associates with V_2. As a consequence, V_1 ends up in penultimate position and gets stressed, and the extra ˈCˈV gets inserted right afterward. Due to open syllable lengthening, an extra pronunciation relation is inserted from the segmental content of V_1 to ˈV. This results in [ˈpjɛgo] 'I fold', where the whole diphthong is as long as a monophthong in the same context (e.g. [ˈpeːlo] 'I peel'; van der Veer 2006). In Fig. 7c we see what happens when /pjɛg/ is concatenated with $/\text{ate}/_{\text{IND.PRS.2PL}}$. In this case, the stress is assigned to V_2, which lengthens. V_1 is unstressed, so it must stay short. Despite this, both /j/ and /ɛ/ are pronounced, and the form surfaces as [pjeˈgaːte] 'you fold', where the whole diphthong is as short as a monopthong in the same context (e.g. [peˈlaːte] 'you peel'; van der Veer 2006).

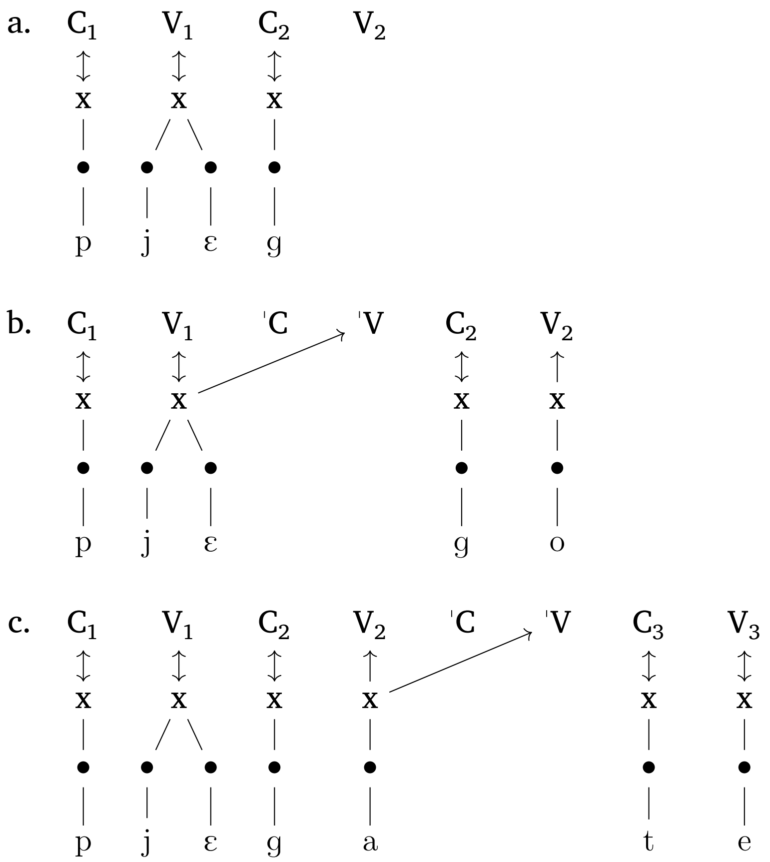

Figure 7: TT non-alternating diphthongs

4 Conclusions

The difference between alternating and non-alternating diphthongs can be related to their phonological representations: in non-alternating diphthongs, two vocoids are two subsegmental components, whereas in alternating diphthongs they are two segments. These structures differ from monophthongs, which contain no glide. Thus, the main argument against a fully phonological account of *dittongo mobile*, namely the fact that not all /jɛ/ and /wɔ/ surface as [e] and [o] in unstressed position, and that not all /e/ and /o/ surface as [jɛ] and [wɔ] in open, stressed syllables, can be rebutted. This allows us to dispense with an allomorphy-based analysis in favor of a fully phonological account. I maintain that the latter is preferable, on the one hand because a system which derives different surface forms from one and the same underlying form reduces the idiosyncrasy of the system, and on the other hand, because it aligns better with speakers' intuitions about the relatedness of different phonetic realizations of paradigmatically equivalent objects.

This analysis also provides further support to the hypothesis that phonetically similar objects (viz the diphthongs and monophthongs discussed in this paper) can conceal different phonological representations, as already maintained e.g. by Lowenstamm (2003) for *mutae cum liquida*, who explicitly argues "for the need to recognise a dual phonological analysis, one monosegmental and one bisegmental, of phonetic segments involving secondary articulation", but also by Dresher & Compton (2011), who distinguish between palatalizing and non-palatalizing [i] in Inuit dialects, Spencer (1986), who does the same for Czech, Gussmann (2007), who distinguishes between three different kinds of [ɛ] in Polish, and Cavirani (2022), who distinguishes between silent and empty nuclei. Crucially, in all these cases, what matters is the phonological behaviour of these objects, rather than their phonetic properties. Thus, this paper also provides some support for the substance-free take on phonology (Blaho 2008; Samuels 2012; Dresher 2014; Iosad 2017; Dresher 2018; Scheer 2019; Chabot 2021; Odden 2022 a.o.)

Finally, an articulated segmental representational format such as that proposed in this paper, where a distinction is made between segmental and C/V nodes, possibly allows for a consistent representation of virtual geminates (Ségéral & Scheer 2001), namely of objects that display the behaviour of geminates but surface as singletons. These objects could indeed be represented as in Fig. 8, where the C nodes projects two segmental nodes, only one of which is pronounced.

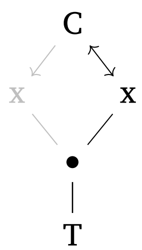

Figure 8: TT virtual geminates

I leave the development of this hypothesis for future research.

References

Anderson, Stephen R. 1976. Nasal consonants and the internal structure of segments. *Language* 52(2). 326–344.

Blaho, Sylvia. 2008. *The syntax of phonology: a radically substance-free approach.* University of Tromsø dissertation.

Booij, Geert & Bart van der Veer. 2015. Allomorphy in OT: the Italian mobile diphthongs. In Eulalia Bonet, Maria-Rosa Lloret & Joan Mascaró (eds.), *Understanding allomorphy: perspectives from Optimality Theory*, 45–69. London: Equinox.

Canalis, Stefano. 2018. The status of Italian glides in the syllable. In Roberto Petrosino, Pietro Cerrone & Harry van der Hulst (eds.), *From Sounds to Structures*, 3–29. Mouton de Gruyter.

Carvalho, Joaquim Brandão de. 2019. Representation vs derivation: the case for a modular view of phonology. *Radical*.

Cavirani, Edoardo. 2015. *Modeling phonologization: vowel reduction and epenthesis in Lunigiana dialects*. LOT Dissertation series dissertation.

Cavirani, Edoardo. 2022. Silent lateral actors: the role of unpronounced nuclei in morpho-phonological analyses. *The Linguistic Review* 39(4).

Cavirani, Edoardo & Marc van Oostendorp. 2017. On silent markedness. In Bridget Samuels (ed.), *Beyond markedness in formal phonology*, 101–120. Amsterdam: Benjamins.

Cavirani, Edoardo & Marc van Oostendorp. 2019. Empty morphemes in Dutch dialect atlases: reducing morphosyntactic variation by refining emptiness typology. *Glossa* 4(1). 1–22.

Cavirani, Edoardo & Marc van Oostendorp. 2020. A theory of the theory of vowels. In Kuniya Nasukawa (ed.), *Morpheme-internal recursion in phonology*, 37–56. Berlin/Boston: Mouton de Gruyter.

Chabot, Alex. 2021. *Possible and impossible languages: naturalness, third factors, and substance-free phonology in the light of crazy rules*. Université Côte d'Azur dissertation.

Davis, Stuart. 1990. Italian onset structure and the distribution of il and lo. *Linguistics* 28. 43–55.

De Castro-Arrazola, Varun, Edoardo Cavirani, Kathrin Linke & Francesc Torres-Tamarit. 2015. A typological study of vowel interactions in Basque. *Isogloss. A journal on variation of Romance and Iberian languages* 1(2). 147–177. `https://doi.org/https://doi.org/10.5565/rev/isogloss.19`.

Dresher, Elan. 2014. The arch not the stone: universal feature theory without universal features. *Nordlyd* 41. 165–181.

Dresher, Elan. 2018. Contrastive hierarchy theory and the nature of features. In *Proceedings of the 35th West Coast conference on formal linguistics*, vol. 35, 18–29.

Dresher, Elan & Richard Compton. 2011. Palatalization and "strong i" across Inuit dialects. *Canadian Journal of Linguistics* 56. 203–228.

Egré, Paul. 2015. Explanation in linguistics. *Philosophy Compass* 10(7). 451–462.

Faust, Noam, Nicola Lampitelli & Shanti Ulfsbjorninn. 2018. Articles of Italian unite! Italian definite articles without allomorphy. *Canadian Journal of Linguistics* 63(3). 1–27.

Faust, Noam & Francesc Torres-Tamarit. 2017. Stress and final /n/ deletion in Catalan: combining strict CV and OT. *Glossa* 2(1).

Garvin, Karee, Myriam Lapierre & Sharon Inkelas. 2018. A Q-theoretic approach to distinctive subsegmental timing. In *Annual meeting of the Linguistic Society of America*, 1–13.

Goldrick, Matthew. 2001. Turbid output representations and the unity of opacity. In Masako Hirotani, Andries Coetzee, Nancy Hall & Ji-yung Kim (eds.), *NELS 30*, 231–246. Amherst: GLSA.

Gussmann, Edmund. 2007. *The phonology of Polish*. Oxford: Oxford University Press.

Harris, John & Edmund Gussmann. 2002. Word-final onsets. *UCL Working Papers in Linguistics* 14. 1–42.

Iosad, Pavel. 2017. *A substance-free framework for phonology: an analysis of the Breton dialect of Bothoa*. Edinburgh: Edinburgh University Press.

Kaye, Jonathan. 1981. Les diphtongues cachées du Vata. *Studies in African Linguistics* 12. 225–244.

Kaye, Jonathan. 1985. On the syllable structure of some West African languages. In D. Goyvaerts (ed.), *African Linguistics: Essays in memory of M.W.K. Semikanke*, 285–308. Amsterdam: John Benjamins.

Krämer, Martin. 2009. *The phonology of Italian*. Oxford: OUP.

Krämer, Martin. 2022. Italian. In Christoph Gabriel, Randall Gess & Trudel Meisenburg (eds.), *Manual of Romance phonetics and phonology*, 559–596. Mouton de Gruyter.

Lampitelli, Nicola. 2010. Nounness, gender, class and syntactic structures in Italian nouns. In Reineke Bok-Bennema, Brigitte Kampers-Manhe & Bart Hollebrandse (eds.), *Romance languages and linguistic theory 2008. Selected papers from 'Going Romance' Groningen 2008*, 195–214. Amsterdam: Benjamins.

Lampitelli, Nicola. 2017. A morphophonological analysis of the velar insert in Italian verbs. *Glossa* 2(1).

Larsen, Bergeton. 1998. Vowel length, raddoppiamento sintattico and the selection of the definite article in Italian. In Patrick Sauzet (ed.), *Langues et grammaire II–III, Phonologie*, 87–102. Paris: Université Paris 8.

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, Manchester: 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/Boston: Mouton de Gruyter.

Marotta, Giovanna. 1988. The Italian diphthongs and the autosegmental framework. In Pier Marco Bertinetto & Michele Loporcaro (eds.), *Certamen phonologicum*, 389–420. Torino: Rosenberg & Sellier.

Marotta, Giovanna. 1993. Selezione dell'articolo e sillaba in italiano: un'interzione totale? *Archivio Glottologico Italiano* 15. 255–296.

McCrary, Kristie Marie. 2004. *Reassessing the role of the syllable in Italian phonology: an experimental study of consonant cluster syllabification, definite article allomorphy and segment duration*. Los Angeles: University of California dissertation.

Muljačić, Žarko. 1971. Gli allofoni /il/, /lo/ e /l/ e la fonologia jakobsoniana. *Lingua Nostra* 32. 82–84.

Muljačić, Žarko. 1974. Ancora sulla forma dell'articolo determinativo italiano. *Italica* 51. 68–71.

Odden, David. 2022. Radical substance-free phonology and feature learning. *Canadian Journal of Linguistics* 67(3).

Passino, Diana. 2009. An element-based analysis of Italian nominal inflection. In Montermini, Boyé & Tseng (eds.), *Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux*, 63–75. Somerville, Massachusetts: Cascadilla Press.

Passino, Diana, Joaquim de Carvalho & Tobias Scheer. 2022. Syllable structure and (re)syllabification. In Christoph Gabriel, Randall Gess & Trudel Meisenburg (eds.), *Manual of Romance phonetics and phonology*, 89–126. Mouton de Gruyter.

Petrosino, Roberto. 2018. Allomorphy of Italian determiners at the morphology-phonology interface. In Roberto Petrosino, Pietro Cerrone & Harry van der Hulst (eds.), *From Sounds to Structures*, 295–328. Mouton de Gruyter.

Polgárdi, Krisztina. 1999. Constraint ranking, government licensing and the fate of final empty nuclei. In John Rennison & Klaus Kühnhammer (eds.), *Phonologica 1996*, 167–182. The Hagues: Holland Academic Graphics.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: constraint interaction in Generative Grammar*. Blackwell.

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

Repetti, Lori. 2020. The masculine singular definite article in Italian: the role of the syllable. *Italian Journal of Linguistics* 32(2). 209–232.

Rowicka, Grażyna. 1999. *On Ghost Vowels: A Strict CV Approach*. The Hague: Holland Academic Graphics.

Russi, Cinzia. 2006. A usage-based analysis of the allomorphy of the Italian masculine definite article. *Studies in Language* 30(3). 575–598.

Saltarelli, Mario. 1970. *A phonology of Italian in a Generative Grammar*. The Hague: Mouton de Gruyter.
Samuels, Bridget. 2012. The emergence of phonological forms. In Anna Maria Di Sciullo (ed.), *Towards a biolinguistic understanding of grammar. Essays on interfaces*, 193–213. Benjamins.
Scheer, Tobias. 2004. *A lateral theory of phonology*. Vol. 1. Berlin: Mouton de Gruyter.
Scheer, Tobias. 2012. Invariant syllable skeleton, complex segments and word edges. *Journal of Linguistics* 48. 685–726.
Scheer, Tobias. 2017. Melody-free syntax and phonologically conditioned allomorphy. *Morphology* 26. 341–378.
Scheer, Tobias. 2019. Phonetic arbitrariness: a cartography. *Phonological Studies* 22. 105–118.
Scheer, Tobias. 2022. 3 x Phonology. *Canadian Journal of Linguistics* 67(3). 1–56.
Scheer, Tobias & Eugeniusz Cyran. 2018. Syllable structure in Government Phonology. In S. J. Hannahs & Anna R. K. Bosch (eds.), *The Routledge handbook of phonological theory*, 262–292. London & New York: Routledge.
Ségéral, Philippe & Tobias Scheer. 2001. Abstractness in phonology: the case of virtual geminates. In Katarzyna Dziubalska-Kołaczyk (ed.), *Constraints and preferences*, 311–337. Mouton de Gruyter.
Shih, Stephanie & Sharon Inkelas. 2019. Autosegmental Aims in Surface-Optimizing Phonology. *Linguistic Inquiry* 50(1). 137–196.
Sluyters, Willibrord. 1992. *Representing diphthongs*. University of Nijmegen dissertation.
Spencer, Andrew. 1986. A non-linear analysis of vowel-zero alternations in Polish. *Journal of Linguistics* 22. 249–280.
Torres-Tamarit, Francesc. 2015. Length and voicing in Friulian and Milanese. *Natural Language & Linguistic Theory* 33(4). 1351–1386.
Tranel, Bernard & Francesca Del Gobbo. 2002. Ocal conjunction in Italian and French phonology. In Caroline Wiltshire & Joaquim Camps (eds.), *Romance phonology and variation*, 191–218. Amsterdam: John Benjamins.
Van der Veer, Bart. 2006. *The Italian 'Mobile Diphthongs'. A test case for experimental phonetics and phonological theory* (Ph.D. dissertation). Utrecht: LOT.
Van Oostendorp, Marc. 2008. Incomplete devoicing in formal phonology. *Lingua* 118(9). 1362–1374.
Zimmermann, Eva. 2017. The power of a single representation: Morphological tone and allomorphy. *Morphology* 26. 269–294.