

A split-gesture, competitive, coupled oscillator model of syllable structure predicts the emergence of edge gemination and degemination

Francesco Burroni
Department of Linguistics
Cornell University
fb279@cornell.edu

Abstract

The phonological mechanisms responsible for the emergence of edge geminates in phonological processes like the Italian *Raddoppiamento (Fono-)Sintattico* (RS) are an open issue. Previous analyses of Italian treat gemination of (i) word initial consonants, (ii) morpheme-final consonants, and (iii) word final consonants as separate processes brought about by dedicated rule/constraints. We argue that these edge gemination processes result from the same, independently established principles. Through computational simulation of the split-gesture, competitive, coupled oscillator model of syllable structure of Articulatory Phonology, we show that increases in closure duration typical of geminates arise from changes to consonant/vowel couplings. Word initial gemination follows from coupling of a closure gesture to a preceding vowel across a word boundary. Word final gemination follows from coupling of a release gesture to a following vowel. In both cases, the posited structures reflect changes in syllabification hypothesized in previous work. The model simulation also predict different durations for resyllabified edge geminates and medial lexical geminates, in line with experimental findings on the topic. Changes to consonant/vowel couplings also account for the opposite effect: word initial degemination. Thus, the coupled oscillator model of Articulatory Phonology, originally developed to model intergestural timing, predicts the emergence of edge gemination/degemination.

1 Introduction

Word initial and word final geminates, collectively known as edge geminates, are employed

contrastively in a highly restricted subset of the world's languages (Burroni and Maspong, To appear; Kraehenmann, 2011; Topintzi and Davis, 2017). This limited cross-linguistic distribution is often attributed to poor perceptual recoverability (Blevins, 2004). Despite the disfavorable phonetic characteristics of edge geminates, speakers of some languages productively create them in the speech stream as a result of regular phonological process. A well-known example is the so-called *Raddoppiamento (Fono-)Sintattico (RS)* in Central and Southern Italo-Romance varieties and Standard Italian (Passino, 2013 and references therein).

Edge-consonant gemination is not a unique feature of Italo-Romance. It has also been reported in a variety of typologically diverse and genetically unrelated languages (Bertinetto and Loporcaro, 1988), such as Finnish (Bertinetto, 1985), Biblical Hebrew (Lowenstamm, 1996), Pattani Malay (Paramal, 1991), Somali (Bertinetto and Loporcaro, 1988), Seri (Marlett and Stemberger, 1983), and Tamil (Ramasamy, 2011). Edge gemination is, thus, a phenomenon with clear cross-linguistic status, yet our understanding of it remains limited.

Three issues stand out in the discussion of edge geminates. The first issue is that, even though word initial gemination is by far the most widely studied case, other types of edge gemination also exist. Central and Southern Italian speakers, for instance, geminate initial consonants, as well as morpheme/word final consonants. Unified treatments of the phenomena have, however, rarely been pursued (for an exception *cf.* Passino, 2013; and partly Chierchia, 1986). Accordingly, the relationship among different types of edge gemination, if any, remains unclear.

The second issue is that phonological accounts represent derived initial geminates and medial lexical geminates with identical ambisyllabic

structures (Section 2). Crucially, there are systematic phonetic differences between the two. Edge geminates are consistently shorter than medial geminates, as experimental work on Italian shows (Payne, 2005; Campos-Astorkiza, 2012). These differences in duration are unexpected in current phonological accounts.

The third problem concerns the relationship between the emergence and loss of edge geminates. The emergence of edge geminates in Italian varieties and other languages has been analyzed as the synchronic consequence of regular phonological process. The loss of edge geminates, on the other hand, has been treated as a diachronic process as a consequence of perceptual/articulation biases and exemplar dynamics (Burroni and Maspong, To appear; Blevins and Wedel, 2009; Blevins, 2004). Nevertheless, synchronic degemination has been documented for Swiss German dialects (Kraehenmann and Jaeger, 2003) and synchronic diffusion of degemination has been documented for Pattani Malay (Burroni et al., 2020). Therefore, even though edge gemination and degemination share the basic property of altering consonantal duration, the mechanisms posited to account for them are remarkably different in both their motivation and timescale. No model of the relation between perceptual biases and changes in articulation has been developed either.

We argue that all types of edge gemination processes observed in languages like Central/Southern Italo-Romance varieties and Italian follow from changes to the dynamical coupling of consonants and vowels, which reflect changes in syllabification in a split-gesture, competitive, coupled oscillator model of syllable structure (Nam et al., 2009; Nam, 2007a; Nam, 2007c). This model also predicts the attested differences in duration between derived edge geminates and lexical medial geminates. Finally, changes in dynamical coupling between consonants and vowels also capture edge degemination, thus, providing a unified account of both phenomena.

2 Empirical phenomena under investigation and previous analyses

We investigate two set of empirical phenomena: (i) edge gemination in languages like Central/Southern Italo-Romance varieties and Italian and (ii) word initial degemination in

languages like Swiss German and Pattani Malay. There are three different edge gemination processes in Italian.

First, speakers are known to produce new word initial geminates in the context of *RS*, provided that the target consonant is not already long. A word w_i undergoes *RS* if: (i) the preceding word w_{i-1} is stressed on the final syllable, e.g., /fa'rɔ 'bene/ → [fa'rɔ 'b:ene] 'I will do well' and (ii) w_{i-1} belongs to a closed class of monosyllables or disyllabic forms that do not have final stress but nonetheless trigger *RS*, e.g., /'kome 'maj/ → ['kome 'm:aj] 'how come'. Second, singleton word final codas, usually only present in loanwords, are geminated before a vowel initial suffix in morphological derivatives, e.g., /buldog/ + /-ino/ → [buldog:ino] 'small bulldog'. Third, word final codas are also geminated phrasally preceding another vowel initial word, e.g., /buldog ag:res:ivo/ → [buldog: ag:res:ivo] 'aggressive bulldog', a phenomenon often labeled *backwards RS*. Morpheme/word final gemination is subject to variation for final sonorants, especially [r], but it is categorical for obstruents (Passino, 2013). These three gemination phenomena are rarely offered a unified treatment, as the focus is usually on *RS* alone.

RS, the first type of gemination and the one that is most often treated in phonological work, is also subject to a fair amount of dialectal variation (Loporcaro, 1997), as, in some Italo-Romance varieties or regional pronunciations of Standard Italian, the process is triggered only after a small subset of lexical items or is absent altogether.

There are three main analyses of *RS*. The first approach holds that *RS* is a byproduct of well-formedness conditions on Italian (final) stressed rhymes or metrical feet. Under this approach, *RS* is due to speakers geminating a word initial consonant to create an ambisyllabic geminate. This ambisyllabic geminate makes a final stressed syllable closed and, thus, heavy, in conformity with a requirement that all Italian stressed syllables either have coda or contain a long vowel. A second approach holds that words that trigger *RS* contain an underlying, featurally empty consonantal slot that only surfaces *via* total assimilation in *RS* environments. Insertion of an entire CV skeleton has also been proposed to account for *RS*, and morpheme/word-internal gemination as well (Passino, 2013). A third approach holds that productive *RS* is limited to a post-tonic environment, accordingly, the only rule needed is a

gemination rule of word initial consonants after word ending in stressed vowels.

None of these solutions is unproblematic. Well-formedness conditions on stress rhymes are at odds with the fact that *RS* also takes place after words that do not have final stress and that certain varieties also show no relationship between final stress and *RS* (Loporcaro, 1997). Empty consonant slots never surface and are, thus, problematic from an acquisition perspective. Additionally, there are words that trigger *RS* but did not have final consonants even when we look at the Latin ancestors of these words. Finally, reducing *RS* to a rule of onset gemination after final stressed vowels comes at the cost of greatly reducing the empirical coverage of the analysis, while certain assumptions regarding rule ordering are also necessary to prevent overapplication in contexts where stressed vowels do not trigger *RS*. All analyses agree that *RS* is produced by changes in syllabification, but they disagree on the rationale.

We show in the next sections that in a split-gesture coupled oscillator model of syllable structure all types of gemination follow purely from syllabification principles in a dynamical model, where no additional rationale is needed. We further show that this model predicts the observed phonetic differences between lexical medial geminates and derived edge geminates, a fact that is missed by other accounts.

The second set of phenomena we investigate are degemination processes. Degeminations of initial geminates has been reported after obstruent-final words in Swiss German, e.g., /s t:an̩kxə/ → [s tan̩kxə] ‘the filling-up’ (Kraehenmann and Jaeger, 2003). Degeminations of initial geminates has also been reported for Pattani Malay, as some minimal pairs with and without initial geminates onsets are merging, e.g., [dapo] ‘kitchen’ and [d:apo] ‘at the kitchen’ are often no longer distinguishable in terms of closure duration of the initial consonant (Burroni et al., 2020). Degemination in Swiss German has been attributed to the loss of one of the two timing slots associated with initial geminates after obstruent-final words. The Pattani Malay neutralization has been analyzed in an exemplar model as a random walk in closure duration space leading to merger (Burroni and Maspong, To appear following Blevins and Wedel, 2009). In both cases a poor perceptual recoverability is invoked to drive change in the phonological representation of words, yet no link with the

production of singletons and initial geminates has been explicitly proposed. We show that degemination also follows from changes in coupling reflecting changes in syllabification in a split-gesture, competitive, coupled oscillator model of syllable structure developed in the framework of Articulatory Phonology (Browman and Goldstein, 2000; Nam, 2007a).

3 The Articulatory Phonology split-gesture, competitive, coupled oscillator model of syllable structure

In the framework of Articulatory Phonology (AP) phonological primitives are identified with articulatory gestures. Gestures are conceptualized as time dependent driving forces that modify the value of tract variables (TVs) and the positions of the synergy of articulators associated with TVs. An example of Lip Aperture being driven until time 10 to a value of 0 mm, representing a bilabial closure [b], and until time 20 back to its original starting value of 10 mm, representing the release of the closure, is presented in Figure 1.

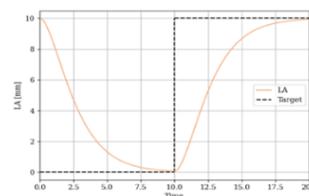


Figure 1 Example of Lip Aperture (LA) constriction and release, implemented with the model in Appendix A.

In the original Task Dynamic model of AP, the duration and relative timing of each gesture was considered part of the lexical representation of words and specified “by hand”. Browman and Goldstein (1990) later modelled the unfolding of a gesture in time with a “virtual” second order undamped systems that has the same stiffness of the original gestural system. The onset and target achievements of the gesture were arbitrarily identified with 0° and 240° of this virtual gestural cycle. Gestures could then be timed to each other by referring to phase relationships of their virtual cycles, e.g., synchronously (0° to 0°) or onset to target of the preceding gesture (0° to 240°). Other phase relationships were deemed possible, but the number of linguistically relevant ones was hypothesized to be highly constrained. Intergestural timing was modeled under a working

assumption that, for n -gestures, at maximum $n - 1$ local coupling forces between gestural pairs could be specified. All relative timing relationships could thus be defined in terms of coupling to a preceding gesture (Browman and Goldstein, 2000).

A more principled dynamical model of relative timing between two gestures was developed by Saltzman and Byrd (2000) using coupled oscillators. Saltzman and Byrd (2000) showed that punctate relative phases (or ranges of relative phases) can be generated by coupling the oscillators regulating the virtual gestural evolution cycles. The relative phase of two gestures is defined as the difference of their phases (ϕ) around the virtual unit cycle, i.e., $\psi = \phi_2 - \phi_1$. The relative phase task-space potential employed by Saltzman and Byrd (2000) is a simple cosine function:

$$V(\psi) = -a \cos(\psi - \psi_0) \quad (1)$$

In this model, a represent a parameter that controls how quickly target relative phase is achieved. ψ represents the current relative phase value of the system. ψ_0 represent a target relative phase. From this potential function a coupling force, defined as the negative of the derivative of the potential function is derived:

$$-\frac{dV(\psi)}{dt} = -a \sin(\psi - \psi_0) \quad (2)$$

The force function is added to each hybrid oscillator's equation to ensure that the coupled oscillators achieve the relative phase specified at the bottom of the potential valley and complete phase-locking, Appendix B. The coupled oscillator model developed by Saltzman and Byrd (2000) was extended to constellations larger than two gestures by Nam and Saltzman (2003), who challenged the assumption that gestures are timed locally to the preceding gesture. Following Browman and Goldstein (2000), Nam and Saltzman (2003) introduced the possibility of competitive coupling: several, mutually incompatible relative phase targets could now be specified for each pair of gestures. The consequence of competitive coupling is that surface relative timing among different gestures is a "compromise" of different relative phase equilibria specified for coupled oscillators. Nam and Saltzman (2003) focused on c-center timing, a non-local timing regime where the initiation of a word initial vowel gesture appears to be timed with

the midpoint of the onset consonants forming a cluster. Nam and Saltzman (2003) showed that c-center, problematic for strictly local timing as it involves timing to an entire cluster, emerges spontaneously if competing relative phases are specified between two onset consonants and for each consonant to the vowel.

A full competitive model of syllable structure in Articulatory Phonology was developed by Nam (2007a; 2007b; 2007c). Nam proposed that the articulatory gestures associated with syllables can be represented as nodes in an undirected graph with no loops, where edges represent target phase couplings for the gestural nodes they connect. Using this graph representation, competing target relative phases can be specified for each gestural pair and competitive coupling is generalized to all possible gestural pairs.

A second feature of the model is that consonantal gestures were split into two gestures: a closure and a release gesture. Nam (2007b), following Browman (1994), argued that releases should be treated as separate gestures, rather than as a return to a neutral vocal tract position. The reason is that the stiffness and velocity of closures and releases are similar, thus, suggesting that both are actively controlled gestures. Nam (2007a; 2007c) also showed that vowels can display c-center timing to the midpoint of the closure and release of a single consonant onset, Figure 2. This another fact that can be taken as evidence for a multigestural representation of a single consonant, similar to that of clusters.

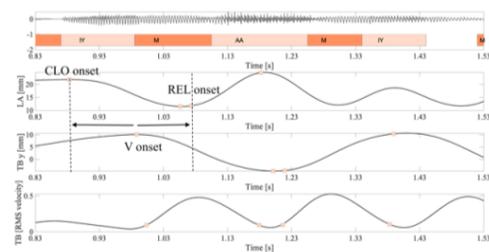


Figure 2 Electromagnetic articulography data exemplifying single consonant c-center for an English speaker producing the word *mommy*. Vowel onset is symmetrically displaced between closure and release.

C-center in singleton consonants has since been experimentally confirmed and further studied (Tilsen, 2017). Nam (2007a) also showed that many properties of phonological systems can be understood in a split-gesture model; among these

are onset/coda asymmetries, both typological and developmental ones, and moraic structure and its acoustic reflexes. Notably, Nam (2007a) also hypothesized that many properties of geminates are best understood through the lenses of a split-gesture model.

4 The model

The model we present closely follows the one developed by Nam (2007a; 2007c). Syllables are represented as nodes in an undirected graph without loops. This graph is known as the coupling graph. Closure and release belonging to the same consonant are represented as separate nodes. An example of this split-gesture graph representation of CV and VC syllables is illustrated in Figure 3.

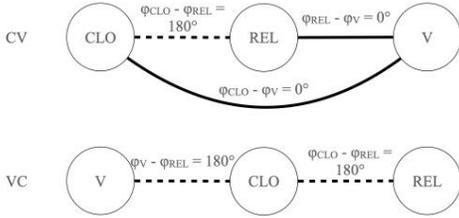


Figure 3 Split-gesture graph representation of CV (top) and VC (bottom) syllables, dashed lines represent anti-phase coupling, solid line in-phase

Figure 3 showcases another important constraint imposed on the model: only two target relative phases are assumed to be available: 0° and 180° (Tilsen, 2018). These are termed in-phase and anti-phase.

The rationale for only two phases is that only those are readily observed in the realm of human (and animal) movement, e.g., in transitions between different gaits of quadrupeds, like horses. Those two phase relationships have also been shown to emerge in experimental tasks involving rhythmic movement (Turvey, 1990). Other relative phase configurations can only be learned with training or emerge from competitive coupling (Nam, 2007a). In this model, the virtual cycle controlling the timing of each gesture is represented only in terms of phase around the unit circle. The differential equation controlling the evolution of each oscillator’s phase in the system of coupled oscillators is defined as:

$$\dot{\theta}_i = \omega_i + \sum_{j=1}^N K \sin(\theta_j - \theta_i - \psi_0) \quad (3)$$

ω_i is the natural frequency of the i^{th} oscillator, set to 2π for our simulations. K is a coupling constant that determines the force exerted by each pair in settling towards target relative phase

equilibria. θ_j with $j = 1 \dots n$ is the j^{th} oscillator’s phase to which θ_i is coupled. ψ_0 is a relative phase target equilibrium for the relative phase of θ_i and θ_j . The model generalized Saltzman and Byrd’s (2000) model to a larger system of oscillators. The matrix form of the model is presented in Appendix C.

This model returns $\dot{\theta}$, an $i \times 1$ vector of oscillator phases at each time step of the simulation of the differential equation. All differential equations were numerically integrated in MATLAB using a forward Euler method over a time range $[0 \ 100]$, the time step was fixed at $.1$. Following previous work (Nam, 2007a; Tilsen, 2018), the phase of each oscillator is mapped to a virtual gestural cycle using a cosine function. Gestures are hypothesized to be triggered once phase-locking is completed and the virtual cycle oscillator crosses 0° .

Following Tilsen (2018), we impose a constraint on initial phases such that each gestural oscillator has a higher initial phase than the gestural oscillator following it in the linearly ordered phonological sequence. For instance, for a CV sequence, with C split into CLO-REL, we impose a constraint $\varphi_{CLO} > \varphi_{REL} > \varphi_V$. These constraints on initial phase values are taken to be part of the lexical representation and to reflect learned order of movements (Tilsen 2018).

5 Experiments

5.1 Singleton c-center and geminate timing

The model can generate a variety of previously reported (relative) timing patterns.

Simulation of c-center timing is achieved by coupling the closure (CLO) and release (REL) oscillators with a target relative phase of 180° (anti-phase), while both CLO/vowel (V) and REL/V are coupled with a target relative phase of 0° (in-phase). The results are the stable relative phase patterns displayed in Figure 4 top and middle. CLO and REL have a relative phase of 120° , while CLO and V and REL and V have a relative phase of 60° and -60° respectively.

The model, thus, predicts a symmetric initiation of the V gesture after the initiation of CLO and before the initiation of REL, Figure 4 bottom. Arrows depict the initiation of each gesture after oscillators have settled in stable relative phases.

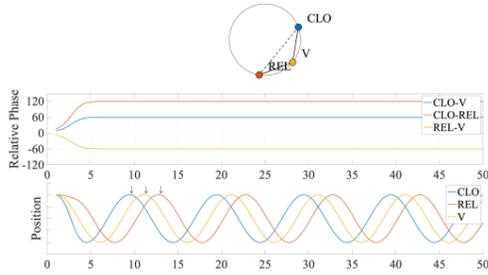


Figure 4 Simulation of single consonant c-center.

As is well-known, lexical geminates differ from singletons in terms of a longer closure duration (Ladefoged and Maddieson, 1996). In a split-gesture model geminates are represented by an increased relative timing between the initiation of the CLO and REL gestures of the same consonant. Since the CLO and REL of a consonant control the same TV, a later initiation of the REL means that CLO will have control of the articulators for a longer period of time. Nam (2007a) suggested that this can be achieved by assuming that the CLO and REL oscillators are anti-phase coupled, like in a singleton, but, crucially, only the REL oscillator is in-phase coupled to the V oscillator. The result of this coupling is complete anti-phase between CLO and REL/V oscillators. This relative phase pattern predicts a longer delay in the initiation of the REL compared to the initiation of CLO, consistent with the longer durations of geminates, Figure 5.

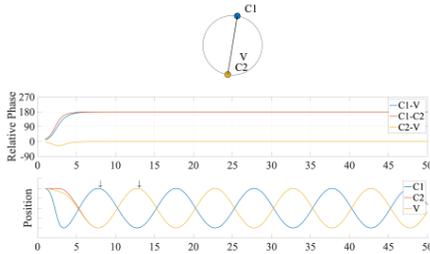


Figure 5 Simulation of geminate timing.

Recent experimental work has shown that the relative timing of closure and vowel initiation for medial geminates is stable across different speech rates (Tilsen and Hermes, 2020). This suggests a stable timing relationship, i.e., in-phase, between the two. Accordingly, a better representation for geminates in a split-gesture model may be coupling only CLO to V, while maintaining anti-phase coupling for CLO and REL. Under this coupling, the result is CLO and V stabilizing in-phase to each other and in anti-phase with REL, Figure 6.

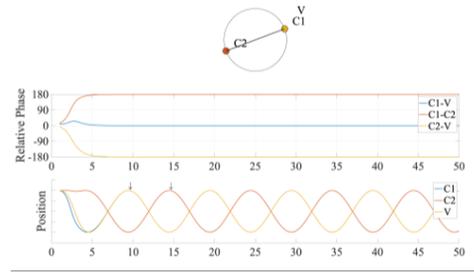


Figure 6 Simulation of geminate timing revised

5.2 Word initial Geminatation

Following previous work (Section 2), we subscribe to the idea that *RS* is a change in syllabification. In particular, *RS* is the formation of an ambisyllabic geminate that acts as both a coda of the preceding syllable and as a word initial onset, as envisioned in all previous analyses. No further dedicated mechanism is necessary for the emergence of word initial geminates. The creation of an ambisyllabic geminate is conceptualized in dynamical terms as follows. The emergence of a new coda amounts to coupling the oscillators of a word final V and a word initial CLO gesture in anti-phase and to decoupling the CLO oscillator from the following V oscillator. No change ensues between the coupling of the CLO and REL oscillators of the word initial consonant, as they still have a target anti-phase relationship. We also assume that the final vowel of the word triggering *RS* and the first vowel of the word undergoing *RS* are anti-phase coupled or sequential. The coupling graph is illustrated in Figure 7.

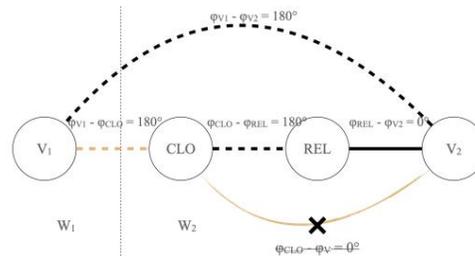


Figure 7 Proposed coupling graph for *RS*.

If we implement these target relative phases in the model, the result is the achievement of a target relative phases between CLO and REL of 135° . This relative phase relationship ensures that the CLO has active control of the TV for a period that is longer than for singleton (120°), but shorter than for lexical geminates (180°), in line with findings showing that *RS* derived edge geminates closure

duration is not as long as that of lexical medial geminates. We return to this issue in Section 6. The model further predicts the correct relative timing initiation: final V of the word triggering *RS*, followed by CLO, followed by REL, followed by V2 of word undergoing *RS*, Figure 8.

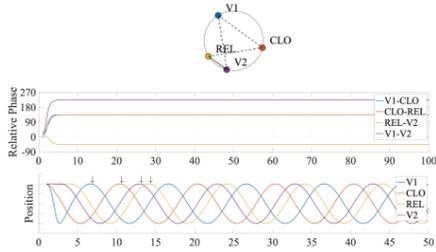


Figure 8 Simulation of *RS*.

5.3 Word final gemination across morpheme/word boundaries

Following previous work detailed in Section 2, especially Passino (2013), we assume that word final gemination across morpheme/word boundaries follows from changes in syllabification, just like *RS*. In this case, the morpheme or word final consonant of the word or stem triggering gemination becomes ambisyllabic and hence geminates, like in *RS*. In dynamical terms, a coupling relationship between the oscillators of a word final REL and a word initial V2 gesture emerges, while the REL oscillator is no longer coupled to the preceding V1 oscillator, as is usually the case for codas that share a mora with preceding vowels and shorten them (Nam, 2007c). No change ensues for the coupling of the CLO and REL oscillators of the word final consonant. They still have a target anti-phase relationship. The coupling graph is illustrated in Figure 9.

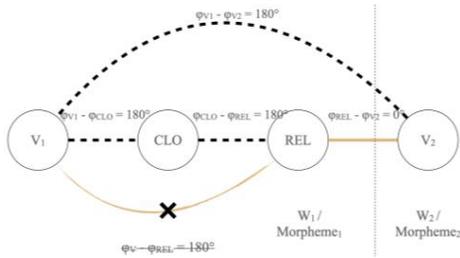


Figure 9 Proposed coupling graph for word/morpheme final gemination.

Exactly as for *RS*, the model predicts a target relative phase between CLO and REL of 135° , Figure 10. The model, thus, generates both the correct relative timing pattern and it also predicts

word final gemination across morpheme/word boundary. Again, derived edge geminates are expected to be shorter than lexical medial geminates.

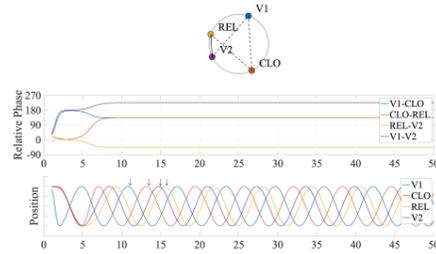


Figure 10 Simulation of word/morpheme final gemination.

5.4 Experiment 4: word initial degemination

In languages like Swiss German and Pattani Malay, synchronic or lexically diffusing degemination of word initial geminates has been observed in experimental work, Section 2. These have been attributed to poor perceptual recoverability triggering changes in phonological representation or in exemplar dynamics of closure duration. Yet, the relationship between degemination and articulation has been left unaddressed. In a split-gesture, competitive, coupled oscillator model of syllable structure incipient degemination can be captured simply as the emergence of a more stable structure where both CLO and REL are in phase coupled to V. Lexical initial geminates are represented with a coupling graph identical to lexical medial geminates: in-phase CLO-V and antiphase CLO-REL. The change in coupling graph structure that triggers degemination is the emergence of a stable in-phase coupling between REL and V, Figure 11.

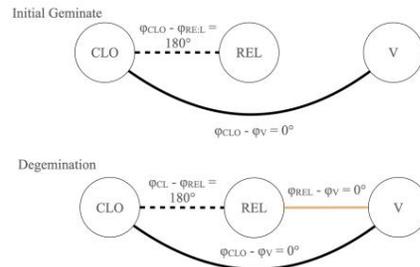


Figure 11 Coupling graphs for initial geminates and word initial degemination.

Obviously, if this coupling graph is used as input to the model, c-center timing emerges. The result is a relative phase between CLO and REL of 120° , identical to that of singletons. Thus, the result of

this change in coupling structure is degemination, as suggested by Nam (2007a), Figure 12.

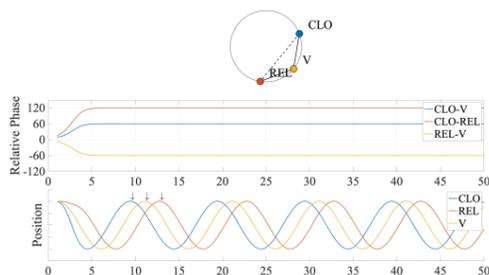


Figure 12 Simulation of word-initial degemination.

5.5 Dynamics of syllabification as the main force behind edge (de)gemination

Having illustrated how the model predicts the emergence and loss of edge geminates, we can now fully appreciate the rationale behind these phenomena: dynamical principles of syllabification in a competitive, coupled-oscillator model of intergestural timing.

The gemination phenomena we have discussed follow only from translating previously hypothesized changes in syllabification into changes to coupling graphs of articulatory gestures forming consecutive syllables. *RS* and morpheme/word final gemination had already been hypothesized to result from the phonological requirement of creating an ambisyllabic geminate; either to create heavy syllables or because of assimilation, originally due to empty consonants (Section 2). In the spirit of dynamical system theory, the model we have presented does not force to choose between these competing alternatives. Instead, the gemination has no further rationale: the process follows purely from the emergence of new dynamical couplings among articulatory gesture. These changes reflect resyllabification near a word boundary, where coupling strengths have long been hypothesized to be weak and gestural sliding has been observed (Browman and Goldstein, 2000).

We can, thus, hypothesize that resyllabification emerges in speech production as a consequence of various factors, e.g., fluctuations in coupling strength due to noise in the production system or because of the effects of speech rate. Once resyllabification alters the dynamical couplings, edge gemination is the natural response of the phonological system. No dedicated rule of edge gemination is needed.

The case of edge degemination follows from slightly different principles. It is not a case of resyllabification across a morpheme/word boundary, but, rather, it represents the emergence of a less marked coupling graph. In other words, it represents a more stable syllabic configuration. Specifically, the emergence of a new coupling between CLO and V, that triggers edge degemination (Figure 11 Coupling graphs for initial geminates and word initial degemination.), represents the emergence of a coupling graph where both articulatory gestures forming a consonant are timed to the vowel. Such configurations with a higher number of links, together with the emergence of in-phase relationships, have been demonstrated to lead to syllable productions that are less sensitive to the effects of noise (Nam, 2007a).

In sum, the model we have presented shows that the emergence and loss of edge geminates are tightly linked as the byproduct of changes to coupling graphs that reflect resyllabification and more stable syllabic configurations.

6 Discussion

We have demonstrated that changes in dynamical couplings, reflecting syllabification, can be responsible for the emergence of (i) word initial gemination, (ii) word/morpheme final gemination, and (iii) word initial degemination. The changes in syllabification were implemented by introducing changes in the dynamical coupling between the oscillator controlling the relative timing of CLO, REL, and V in a split-gesture, competitive, coupled oscillator model of syllable structure. This model offers a unified theory of the articulatory features that accompany the emergence and disappearance of edge geminates.

Furthermore, the model also predicts durational difference between derived edge geminates and lexical medial geminates. This is accomplished by different phase locking patterns: for lexical (medial) geminates the CLO and REL oscillators stabilize at a relative phase of 180° ; for derived edge geminates the relative phase is 135° . Recall that the difference between singleton, displaying c-center timing, and geminates is one of 120° vs 180° . Accordingly, edge geminates only cover $\frac{1}{4}$ ($15^\circ/60^\circ$) of the relative phase difference that separates singleton from geminates. This relative phase patterns are compatible with the experimental findings of Campos-Astorkiza

(2012), who reported that geminates derived via *RS* have a percentage of lengthening, compared to singleton, in the range of 23-60% (on average around 50%). For lexical geminates the range is 200-276%. The model presented, thus, offers not only unified treatment of different types of edge gemination and degemination, but it also predicts phonetic differences between derived initial and lexical medial geminates that align with experimental findings. Crucially, the model does not require any dedicated mechanism to accomplish this, the phonological processes follow purely from dynamical couplings that reflect changes in syllabification. In this way, shared intuitions presented in previous work can be unified without a need for choosing any one rationale, as the system is self-organizing.

The model also has some limitations. First, it accounts for difference between singleton and geminates purely in terms of relative intergestural timing. However, differences between singletons and geminate are likely to be manifested also in intragestural timing due to differences in parameters like targets, stiffnesses, *etc.* Furthermore, translating relative timing into periods of gestural activation intervals is a non-trivial problem, for which a variety of solutions have been proposed (Tilsen, 2018).

A second limitation is that recent experimental evidence (Tilsen and Hermes, 2020) has shown that the onset of geminate release, with respect to either the onset of the closure or the vowel, is linearly delayed as speech rate increases. For singletons the relative timing patterns are relatively unaffected. Tilsen and Hermes (2020) interpreted these different timing regimes as evidence that singletons can be modelled with coupled oscillators, but competitive queuing and feedback based gestural suppressions (Tilsen, 2016) may be necessary to generate the geminate timing patterns.

This is a more general problem of the coupled oscillator model and of the TD model that regulates gestural evolution. They are feedforward systems with no feedback. This assumption is clearly problematic for speech (Shaw and Chen, 2019; Tilsen, 2016; Parrell et al., 2019). Accordingly, scholars have proposed extensions of the model that take feedback into account (Tilsen, 2016; Parrell et al., 2019). Integrating feedback mechanisms for different types of geminates is a direction that needs to be further explored.

The coupled oscillator model is also sensitive to the initial conditions of the simulation. Specifically, it is sensitive to the initial phase of each gestural oscillator. To side step this problem, we have imposed constraints on initial phases that we take to be a reflex of lexical representation and linear ordering (Tilsen, 2018). However, these constraints may betray the need for integrating a competitive queuing model on top of a coupled oscillator model of syllable structure (Tilsen 2016; 2018).

Finally, the coupling structures posited to account for the emergence and disappearance of edge geminates need empirical verification *via* collection of articulatory data, e.g., EMA or real time MRI. Such a dataset may also be a starting point to explore how the creation of new dynamical couplings may emerge in the first place. In particular, we can hypothesize that fluctuations in coupling strength may give rise to trial to trial variability in coupling of consonants at word edges and vowels (Browman and Goldstein, 2000). Ultimately, these changes may be phonologized as changes to coupling graphs. This hypothesis, however, requires empirical testing.

7 Conclusion

We have demonstrated that the AP split-gesture, competitive, coupled oscillator model provides us with a self-organizing model of syllable structure where edge-gemination and degemination emerge from dynamical coupling of closure and release oscillators with vowel oscillators. The model offers a unified analysis of different types of edge gemination and degemination, an aspect that was missing in previous phonological work. Moreover, the model also predicts crucial phonetic differences between derived edge geminates and lexical medial geminates reported in experimental work, but missing in previous phonological analyses. In sum, the coupled oscillator model of Articulatory Phonology, originally designed to model intergestural timing, has proven to be successful at predicting the finer details of elusive phonological processes like edge gemination and degemination.

Acknowledgements

I am indebted to the members of the Cornell Phonetics Lab, Sam Tilsen, and three anonymous reviewers for providing useful feedback on earlier versions of this paper.

References

- Pier Marco Bertinetto. 1985. A proposito di alcuni recenti contributi alla prosodia dell'italiano. *Annali della Scuola normale superiore di Pisa. Classe di lettere e filosofia*, 15(2):581–643.
- Pier Marco Bertinetto and Michele Loporcaro. 1988. On empty segments and how they got that way. In *Certamen phonologicum, Papers from the 1987 Cortona Phonology Meeting*, pages 37–62. Rosenberg and Sellier, Turin.
- Juliette Blevins. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.
- Juliette Blevins and Andrew Wedel. 2009. Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica*, 26(2):143–183.
- Catherine P Browman. 1994. Lip aperture and consonant releases. In *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*, pages 331–353. Cambridge University Press,.
- Catherine P Browman and Louis Goldstein. 2000. Competing constraints on intergestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP. Bulletin de la communication parlée*(5):25–34.
- Francesco Burroni and Sireemas Maspong. To appear. Re-examining Initial Geminates: Typology, Evolutionary Phonology, and Phonetics. In *Historical Linguistics 2019*. John Benjamins.
- Francesco Burroni, Sireemas Maspong, Pittayawat Pittayaporn, and Pimthip Kochaiyaphum. 2020. A new look at Pattani Malay Initial Geminates: a statistical and machine learning approach. In *Proceedings of the 34th Pacific Asia Conference on Language, Information and Computation*, pages 21–29.
- Rebeka Campos-Astorkiza. 2012. Lengthening and prosody in Tuscan Italian. *Anuario del Seminario de Filología Vasca "Julio de Urquijo": International journal of basque linguistics and philology*, 46(1):83–108.
- Gennaro Chierchia. 1986. Length, syllabification, and the phonological cycle in Italian. *Italian Journal of Linguistics*, 8(1):5–34.
- Astrid Kraehenmann. 2011. Initial geminates. *The Blackwell companion to phonology*:1–23.
- Astrid Kraehenmann and Marion Jaeger. 2003. Phrase-initial geminate stops: articulatory evidence for phonological representation. In *Proceedings of the 15th International Conference of the Phonetic Sciences*, pages 2725–2728, Barcelona.
- Peter Ladefoged and Ian Maddieson. 1996. *The sounds of the world's languages*. volume 1012. Blackwell Oxford.
- Michele Loporcaro. 1997. *L'origine del raddoppiamento fonosintattico: saggio di fonologia diacronica romanza*. volume 115. Francke A. Verlag.
- Jean Lowenstamm. 1996. The beginning of the word. In *Phonologica 1996 Syllables!?: Proceedings of the 8th International Phonology Meeting*, pages 153–166, The Hague. Holland Academic Graphics.
- Stephen A Marlett and Joseph Paul Stemberger. 1983. Empty consonants in Seri. *Linguistic Inquiry*, 14(4):617–639.
- Doris Mücke, Anne Hermes, and Sam Tilsen. 2020. Incongruencies between phonological theory and phonetic measurement. *Phonology*, 37(1):133–170.
- Hosung Nam. 2007a. *A gestural coupling model of syllable structure*. Ph.D. thesis, Yale.
- Hosung Nam. 2007b. Articulatory modeling of consonant release gesture. In *International Congress on Phonetic Sciences XVI*, pages 625–628.
- Hosung Nam. 2007c. Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. In *Laboratory Phonology 9*, pages 483–506, Urbana Champaign. Mouton De Gruyter. publisher: Mouton de Gruyter Berlin.
- Hosung Nam, Louis Goldstein, and Elliot Saltzman. 2009. Self-organization of syllable structure: A coupled oscillator model. *Approaches to phonological complexity*, 16:299–328.
- Hosung Nam and Elliot Saltzman. 2003. A competitive, coupled oscillator model of syllable structure. In *Proceedings of the 15th international congress of phonetic sciences*, volume 1.
- Waemaji Paramal. 1991. *Long consonants in Pattani Malay: The result of word and phrase shortening*. PhD Thesis, Faculty of Graduate Studies, Mahidol University.
- Benjamin Parrell, Vikram Ramanarayanan, Srikantan Nagarajan, and John Houde. 2019. The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLoS computational biology*, 15(9):e1007321.

Diana Passino. 2013. A unified account of consonant gemination in external sandhi in Italian: Raddoppiamento Sintattico and related phenomena. *The Linguistic Review*, 30(2):313–346.

Elinor M Payne. 2005. Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association*, 35(2):153–181.

Mohana Dass Ramasamy. 2011. *Topics in the morphophonology of standard spoken Tamil (sst): An Optimality Theoretic study*. PhD Thesis, Newcastle University.

Elliot Saltzman and Dani Byrd. 2000. Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, 19(4):499–526.

Elliot Saltzman and Kevin Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological psychology*, 1(4):333–382.

Jason A Shaw and Wei-rong Chen. 2019. Spatially conditioned speech timing: Evidence and implications. *Frontiers in psychology*, 10:2726.

Sam Tilsen. 2016. Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics*, 55:53–77.

Sam Tilsen. 2017. Exertive modulation of speech and articulatory phasing. *Journal of Phonetics*, 64:34–50.

Sam Tilsen. 2018. Three mechanisms for modeling articulation: selection, coordination, and intention. *Ithaca, NY: Cornell University*.

Sam Tilsen and Anne Hermes. 2020. Nonlinear effects of speech rate on articulatory timing in singletons and geminates. In *12th International Seminar on Speech Production*.

Nina Topintzi and Stuart Davis. 2017. On the weight of edge geminates. *The phonetics and phonology of geminate consonants*, 2:260–282.

Michael T Turvey. 1990. Coordination. *American psychologist*, 45(8):938.

Appendix A: The Task Dynamic Model

In the Task Dynamic model the state of each TV is represented as a second order critically damped oscillatory system, following the Task Dynamic (TD) approach to motor control in speech (Saltzman and Munhall, 1989)

$$m\ddot{x} + b\dot{x} + k(x - T(t)) = 0$$

m represents the articulator mass. It is usually ignored and set to the unit value. b represents the damping coefficient. Critical damping, $b = 2\sqrt{km}$, is assumed to enforce asymptotic target achievement without oscillations. k represents the stiffness parameter, which determines how quickly the target state of the system is achieved. Higher stiffness corresponds to a quicker target achievement. Finally, x and $T(t)$ represent the current positional value of the system and its target state, respectively.

Appendix B: The hybrid oscillator model of Saltzman and Byrd (2000)

In the original coupled oscillator model of oscillators Saltzman and Byrd (2000) the force function is transformed into a task specific coupling force that drives changes in the acceleration of a hybrid oscillator that arises from the combination of a Van Der Pol and Rayleigh limit cycle

$$\ddot{x} = -\alpha\dot{x} - \beta x^2\dot{x} - \gamma\dot{x}^3 - \omega_0^2 x$$

α represents a linear damping term, while β and γ non-linear van der Pol and Rayleigh damping, respectively. ω_0 represents the oscillator natural frequency.

Appendix C: Matrix implementation of the split-gesture, competitive, coupled oscillator model of syllable structure of Articulatory Phonology

The differential equation controlling the system of oscillators in our model is:

$$\dot{\theta} = \omega + \sum_{j=1}^N K \sin(A_j \circ \Phi_j^T - A_j \circ \Phi_j - \Psi_{0,j})$$

ω is the natural frequency of each oscillator and it is hypothesized to be identical for each oscillator, following previous work (Nam, 2007a). Φ is an $i \times j$ ($i = n, j = n$, where n is the number of oscillators in the system) matrix of initial phases for each oscillator i , with the value repeated across columns. A is an $i \times j$ adjacency matrix such that its element a_{ij} is defined as

1 if the oscillator i is coupled with oscillator j , and 0 otherwise. $\Psi_{0,j}$ is an $i \times j$ matrix of target relative phase where each cell $\psi_{0,ij}$ represents a target relative phase for the oscillator pair θ_i and θ_j . If the oscillators are uncoupled the target relative phase is set to 0. K is a matrix of coupling constants. It is set to a unit matrix in all simulations reported to avoid exploding the parameter space, it could however be used to model cross-linguistic differences (Mücke et al., 2020).