Breaking the Rules: A Tour of Non-Regularity in Turkish Phonology

By

Jae Sıdıka Weller

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Linguistics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2026

Date of Final Oral Examination: 03/24/2026

The dissertation is approved by the following members of the Final Oral Committee:

Nâlân Erbil, Professor, German, Nordic, and Slavic+

William Idsardi, Professor, Linguistics

Tom Purnell, Professor, English

Eric Raimy, Professor, English

Rajiv Rao, Professor, Language Sciences

Acknowledgements

There are a great many people I would like to thank. More, I think, than can feasibly be listed here. So here is a non-exhaustive list of people (and animals) to whom I would like to dedicate this dissertation, in no particular order. All listed here made this work possible in ways large and small.

First, there are the five members that comprise my committee: Professors Nâlân Erbil, William Idsardi, Tom Purnell, Eric Raimy, & Rajiv Rao. Thank you for reading the manuscript of this dissertation and providing feedback, as well as all of your comments along the way.

Then, there are various family members I want to acknowledge: my parents, Steve Weller and Laura Weller; my siblings, Alison Joens, Bonnie Evins, Madeline Weller, Kris Weller, Daphne Buecher, and Cameron Buecher; my grandmother, Sıdıka Sabahat Gülşen; my partner, Taylor Auten; my dog, Smile; and my tarantula, Corduroy. Thank you all for your unwavering support of me, even when I did not believe in myself.

Next, there are numerous faculty members, staff, and students at UW - Madison to whom I owe many thanks. Thank you for all of your support and advice to the academic advisor for my program, Becky Shields, as well as Professors Monica Macaulay and Kelly Wright. In addition, I would like to thank Professor Eric Raimy, my advisor, for going above and beyond in terms of commitment to my success and growth both as a researcher and a person. Thank you to Professor Nâlân Erbil for not only being on my committee, but also answering my numerous strange questions about the Turkish language. Finally, I would like to thank Rachyl Hietpas, my good friend and fellow PhD student; her friendship has kept me sane.

I also want to name Professor Jeremy Steffman of the University of Edinburgh, who was my first research mentor, without whom I never would have embarked on this doctoral

journey. Thank you for seeing my potential and encouraging me to achieve things I had never thought possible.

Lastly, I would like to dedicate this dissertation to you: the reader. This is by far and away the largest piece of work I have ever created, and I wrote it because I love weird linguistics. If you’re reading this, you probably do too. Thank you for engaging with my ideas; I hope that you will find herein a wealth of knowledge and information with which to write your own works about weird linguistics

Table of Contents

1 Introduction

1.1 The General Plan/Goals

The general goal of this dissertation is to provide analyses of phenomena in Turkish phonology which are considered in some way unpredictable or ‘exceptional’. While generalizations about language have been represented in myriad ways, from rules to constraints to autosegmental phonology, data which does not fit neatly into these generalizations are not as well understood. Each of the planned topics of my dissertation has some sort of linguistically ‘ill-behaved’ component to them. I will start with a general review of the literature for each of the four subtopics included in this dissertation (Chapter 2). Then, I will cover Turkish vowel harmony (Chapter 3), which has numerous exceptions. Following the discussion of vowel harmony will be an analysis of a related issue: what has been labeled as epenthetic vowels inserted to break up illicit coda clusters (Chapter 4). However, not all of these ‘illicit’ clusters are actually illicit in Turkish. Fourthly, I will move on to emphatic reduplication, which includes the addition of an unpredictable linker consonant (Chapter 5). Finally, I will talk about plosive voicing alternation in Turkish and representing exceptions to what has been labeled as coda devoicing (Chapter 6). I aim to represent these ‘exceptions’ without needing to make use of arbitrary exception diacritics, to take seriously how these non-regularities might be stored in the phonology of Turkish speakers.

Another goal, which is to be achieved along the way, is to apply the *Events, Features, and Precedence* (EFP) framework (see Idsardi 2022) to a particular language– in this case, Turkish. To date, the most in-depth application of an EFP-type framework was Papillon’s (2020) dissertation. Papillon (2020) used *Precedence Relation Oriented Phonology* (PROP), to represent a number of phonological and morphological phenomena across many of the world’s languages. The Papillon dissertation is strong in that it demonstrates that the framework is useful, in general, for phonology and morphology. However, Papillon understandably had neither the time nor the space to focus on developing the theory in depth for particular languages. This dissertation will seek to fill that gap by applying the EFP model to Turkish phonology. A much more detailed description of the framework and background on relevant works can be found in the literature review (Chapter 2), in section 2.1.

1.2 Linguistic Scope

The scope of this work is outlined briefly here. First, since the concern of this dissertation is phonology, the focus will be on the word level. I recognize that linguists still do not fully agree on what, exactly, constitutes a word, so I will do my best to clarify what I mean. Turkish is an agglutinative language, so a word in the vast majority of cases consists of a root followed by suffixes. There are a few exceptions, which I will list here. There are a very small number of loan prefixes, which I will discuss in my account of vowel harmony. There is also compounding, which I do not analyze in this work. In addition, Turkish emphatic reduplication, one of the topics I will discuss at length, is a partial prefixing reduplication. Aside from these cases, when I am discussing Turkish words, I am referring to a root followed by suffixes. Turkish does also have a number of clitics which are positioned following the root, which I do not discuss, and I make no claim about whether clitics constitute affixes or separate words.

Since Turkish is agglutinative, single words can contain enough information in the suffixes to form a whole sentence. So in some cases, an example word may also constitute a full Turkish sentence. However, the use of such examples does not center on sentence or grammatical structure, but rather the sounds in the word.

At times, the topics in this dissertation will certainly abut onto morphology, as affixes will be discussed extensively in my accounts of vowel harmony and plosive voicing alternation. In addition, the representation of Turkish emphatic reduplication will be discussed at length, and this is certainly a topic which may concern morphologists. I do not aim to make any larger claim here about morphology being a part of or not a part of phonology; my aim is to discuss these topics as they are relevant to *phonology* (e.g., suffixes with disharmonic vowels are certainly relevant to phonology of vowel harmony).

Finally, this dissertation focuses primarily on İstanbul Turkish, which is the prestige dialect of Turkish, and is also the dialect that the vast majority of past phonological literature about Turkish has centered on. In some cases, other dialects will be discussed where data is available and space allows. For example, emphatic reduplication in Azerbaijani Turkish will be briefly summarized in the literature review portion about Turkish emphatic reduplication (section 2.5.4). This focus on İstanbul Turkish is in no way a value judgment on other dialects of Turkish, and certainly should not be taken as an endorsement of prescriptivism (that İstanbul Turkish is in some way 'more correct' than other dialects). I know such prescriptivism exists among some Turkish speakers, and I want to be very clear: İstanbul Turkish is no more correct or more deserving of study than any other Turkish dialect. The cause of my focus on İstanbul Turkish is the product of two pragmatic concerns. Firstly, this dissertation must have a specific focus such that it has a reasonable scope. And secondly, the vast majority of past work and collected data has focused on this dialect. And as this dissertation involves no novel data collection, I must engage with the information available from previous studies, which overwhelmingly concerns the İstanbul Turkish dialect.

1.3 A Note on Writing Conventions

Before discussing Turkish phonology, it would be beneficial to make clear the various spelling conventions for Turkish words used in this work, especially for those who may not be familiar with Turkish orthography. Turkish orthography is fairly transparent, meaning it is typically straightforward to look at the spelling of a Turkish word and understand how it is pronounced, if you are familiar with the writing system. However, I want to make sure that this dissertation can be easily understood by phonologists with a wide variety of backgrounds, who may not be familiar at all with the Turkish language. At the same time, I also want to make sure that the data discussed in this dissertation is at least somewhat accessible to people who might have more familiarity with Turkish orthography than with IPA transcriptions. Also, from my own experience as a Turkish speaker, it is much easier for me to understand and write in Turkish orthography than it is to do so in IPA. So, I have decided on a compromise: where possible, I will provide both the Turkish orthography and IPA transcription for Turkish data. When embedded in prose, this will look like the following example: *şapka* [ʃɑpkɑ] 'hat'. Turkish orthography will be listed first in *italics*, followed by an IPA transcription in either square brackets [] or slashes //, followed by an English translation in single quotes ''. The only place I will *not* provide both orthography and IPA is in the graph representations provided later on in this dissertation. Within the graphs, I will provide only IPA, for legibility reasons, and for clarity, since the graphs are strictly phonological representations.

There is one spelling convention I will make use of from Turkish language teaching and linguistic literature about Turkish, that is not used in typical Turkish orthography. I will use a capital letter where alternations exist in suffixes. For a non-high harmonizing vowel, I use the capital letter *A*. For example, the plural suffix *-lAr* [-lAɾ] has a non-high harmonizing vowel which can surface as *a* [ɑ] or *e* [e]. I will use a capital letter *I* for high harmonising vowels, which can surface as *i ü u ı* [i y u ɯ]. (See the literature review and Chapter 3 for more on vowel harmony). Finally, you will also see capital *D* and *C* used in suffixes to denote *t ~ d* [t] ~ [d] and *c ~ ç* [dʒ] ~ [tʃ] alternation respectively. These will match the voicing of the previous sound. (See the literature review and Chapter 6 for more on plosive voicing alternation.)

I have also provided on the following a list of Turkish orthographic letters along with their IPA transcriptions, for anyone who would like to familiarise themselves with either system. It is adapted from a very similar list provided in Göksel & Kerslake (2006:xxii).

(1) *Turkish orthography with IPA and closest English equivalents*
(see also Göksel & Kerslake 2006:xxii)

Vowels

A a	/ɑ/	'pot'
E e	/e/	'pet' or pat
I ı	/ɯ/	'among' or 'putt'
İ i	/i/	'beat' or 'bit'
O o	/o/	off [2]
Ö ö	/œ/	'bet' but with lip rounding similar to German ö
U u	/u/	'boot' or 'book'
Ü ü	/y/	'bit' with lip rounding similar to German ü

Consonants

b	/b/	'bot'[1]
c	/dʒ/	'jet'
ç	/tʃ/	'chat'
d	/d/	'dog'
f	/f/	'fog'
g	/g/	'gap'
ğ	lengthens preceding vowel, is silent, or /j/ (btw front vowels)	
h	/h/	'hoof'
j	/ʒ/	'massage' or 'visage'
k	/k/	'copy'
l	/l/	'low'
m	/m/	'map'
n	/n/	'nap'
p	/p/	'pod'
r	/ɾ/	'butter'
s	/s/	'sip'
ş	/ʃ/	'ship'
t	/t/	'tip'
v	/v/	'very'
y	/j/	'yell'
z	/z/	'zoo'

[1] For simplicity, I am describing the laryngeal contrast in Turkish plosives as 'voiced' and 'voiceless' and transcribing them as such here. This is typically the convention used in other works describing the sound system of Turkish (see Lewis 1967, Underhill 1976, Clements & Sezer 1982, Inkelas & Orgun 1995, Stachowski 2014, Özçelik 2024, and others). In reality, the phonetic (and phonological) situation is likely more complex than this. This issue will be described in detail in section 2.4 of Chapter 2 and Chapter 6.

[2] I think this vowel may only be similar for English speakers without low-back merger, who pronounce this word like [ɔf]. I have the low-back merger, and pronounce the vowel in 'off' as a low back unrounded vowel. I used this example because it is listed in Göksel & Kerlsake (2006).

2 Literature Review

This section contains an explanation of relevant phonological frameworks, along with necessary background information and previous research related to each of the topics covered in this dissertation.

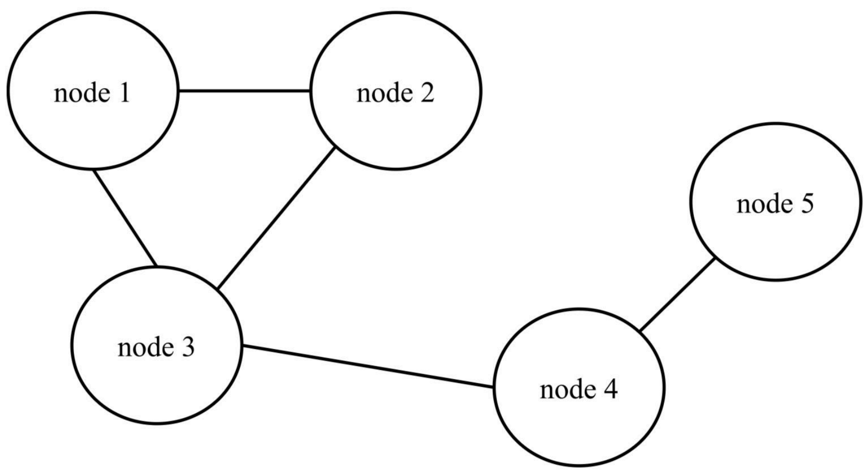
2.1 Events, Features, & Precedence

Phonological representations in this dissertation will make use of the *Events, Features, and Precedence* (EFP) model as laid out in Idsardi (2022). There are two main works which EFP draws from, Raimy (2000)'s multiprecedence and Papillon (2020)'s Precedence Relation Oriented Phonology (PROP), both of which are also laid out later in this section. All of these frameworks use directed graphs to represent phonological phenomena. This section will cover what a graph is, different graph properties and types of graphs. Then, relevant past works on EFP will be summarized.

2.1.1 What is a graph?

A graph consists of a set of vertices or nodes and a set of edges which connect nodes to each other. Graphs can be represented in many ways, but a commonly used and easily human-readable method is to draw the nodes and edges out visually as a diagram. See (2) below for an example of a general graph.

(2) *Example of a graph*

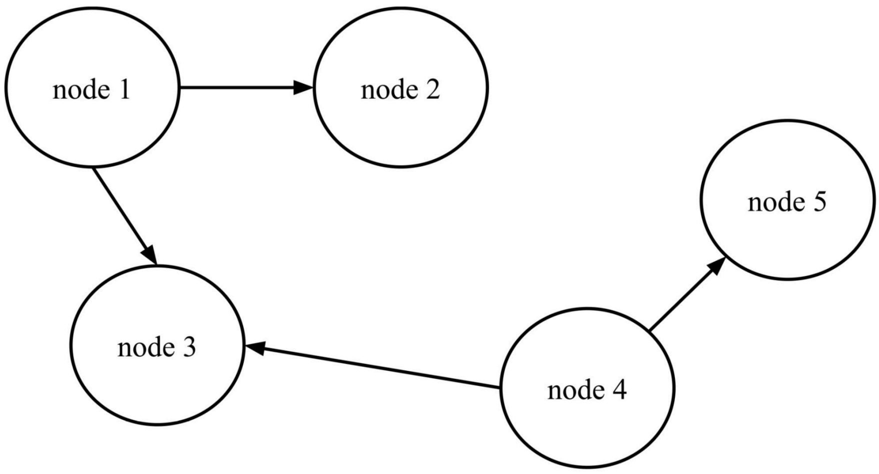


The graph above has five nodes (the circles) and five edges (the lines connecting the nodes). Nodes can be essentially empty points (as above), but they can also hold information, as will be shown later.

The information contained visually in (2) can also be captured using two sets. The first is the set of nodes: {1, 2, 3, 4, 5}. The second is a set of tuples, where each tuple represents one edge of the graph: {(1, 2), (1, 3), (2, 3), (3, 4), (4, 5)}. The numbers in the tuples are nodes which each edge connects; so, for example, the tuple (1, 2) represents the edge connecting node 1 and node 2 in the graph above.

Edges in graphs can be *directed* or *undirected*. An undirected graph, as above in (2), has edges which are all bidirectional; visually these edges are typically depicted with lines. A directed graph (also known as a digraph) has edges with a specific direction, typically depicted as arrows. See (3) below for an example of a directed graph.

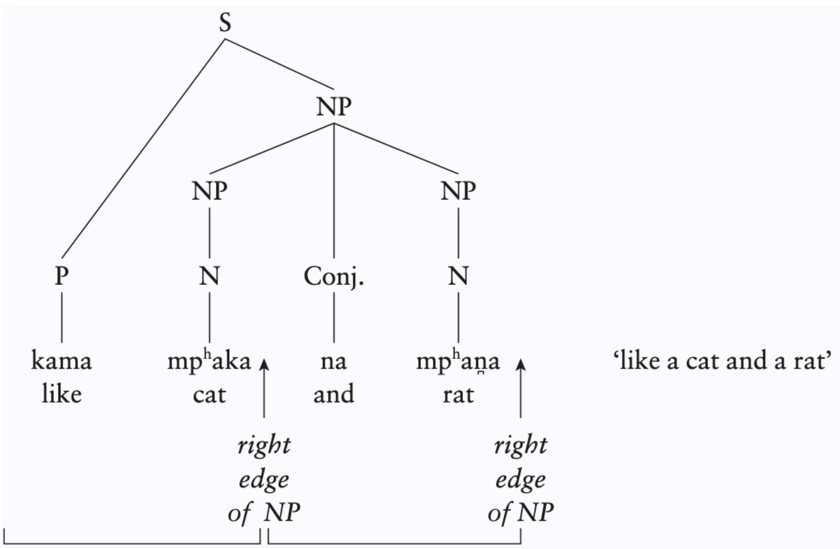
(3) *Example directed graph*



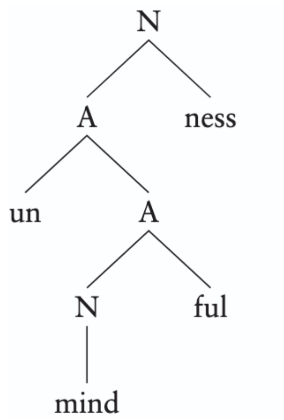
Directed edges capture an asymmetrical relationship between nodes. Exactly what kind of asymmetry depends on what information the graph represents. If the graph in (3) represents a network, for example, perhaps a directed edge represents that a source sends information to another. In that case, the device represented by node 1 sends information to node 2 and node 3. However, node 2 and node 3 do not send information to node 1. In this way, the directed edges represent an asymmetrical relationship in which information flows out from a node, but not back. Another type of asymmetrical relationship which directed edges can represent is a precedence relationship. In this case, nodes represent events in time and a directed edge from one event to another event represents a temporal ordering. So, in the graph above (3), the event represented by node 1 would come before node 2 and node 3 in time. Directed edges represent precedence relationships in EFP, which will be discussed in more detail later in this subsection.

Graphs are a commonly used data structure in linguistics, most frequently seen in the form of trees. Trees are commonly used in syntax (4) to encode sentence structure, in morphology (5) to encode the formation of words from roots and affixes, and in phonology (6) to encode syllable structure. Shown below are examples of each taken from introductory phonology textbooks. Notice that in these language-based examples, the nodes in the graphs contain language information, rather than being empty points.

(4) *Syntax tree from Hayes (2009:216)*



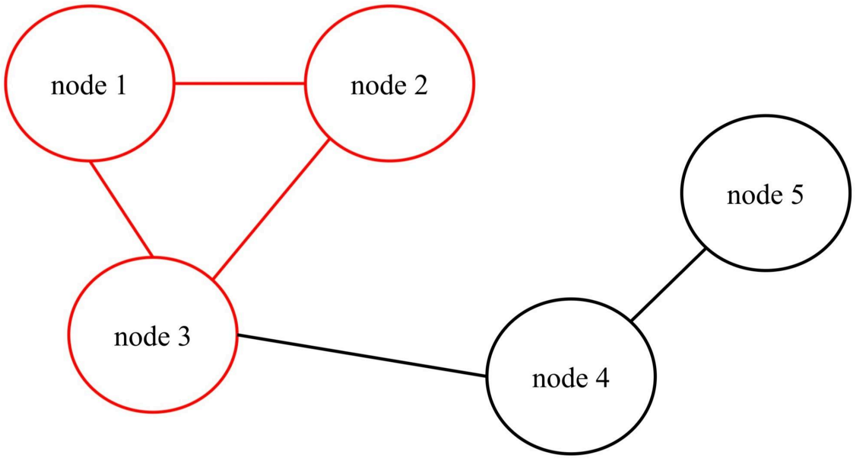
(5) *Morphological tree from Hayes (2009:111)*



(6) *Syllable structure tree from McMahon (2002:110, example 7)*

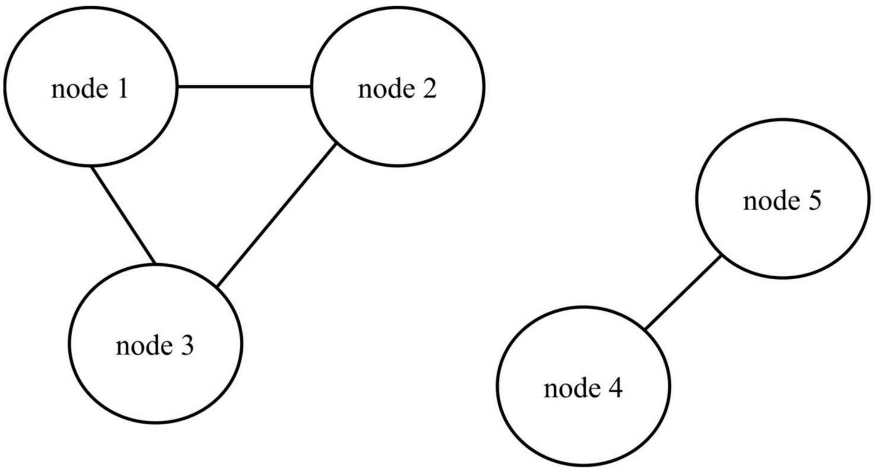
Trees are a restricted type of graph. Specifically, trees are acyclic, connected graphs. Each of these two terms will be defined briefly here. For a graph to be *acyclic*, it must not contain any cycles. A cycle is a path (consisting of one or more unique edges) from a node back to itself. For example, there is a cycle in the graph in (2) from earlier: nodes 1, 2, and 3 form a cycle. This is shown more clearly below in (7) with the cycle highlighted in red. The graph contains a path from node 1 back to node 1: node 1 - node 2 - node 3 - node 1.

(7) *An example of a cycle*



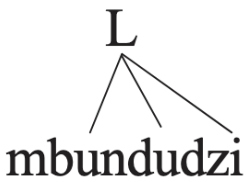
Next, we turn to the requirement of *connectedness*. If a graph is connected, for every pair of nodes in the graph, there exists some path between the two. The example graph in (8) is not connected, since there is no path between nodes {1 2 3} and nodes {4 5}.

(8) *Example unconnected graph*



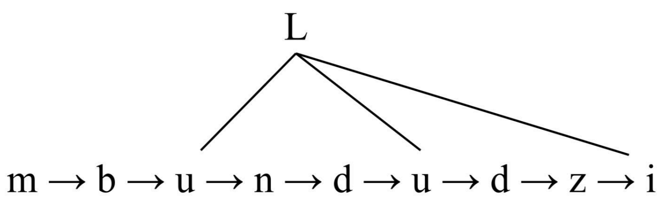
The use of graphs for representations in phonology actually extends beyond trees. The framework of autosegmental phonology (see Goldsmith 1976) makes use of graphs which are not trees. See below for an example from an introductory phonology textbook.

(9) *Tone association lines in Shona from Odden (2005:309, example 25)*



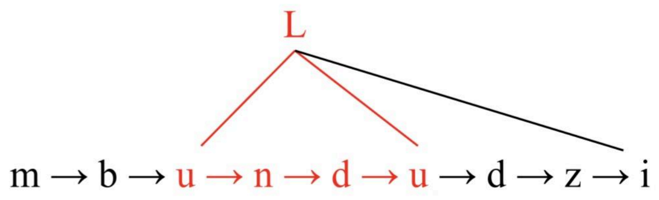
The example in (9) is a graph, but unlike the earlier linguistic examples shown in (4-6), is not a tree. The graph in (9) contains a set of nodes which consist of the phonemes and tones as well as a set of edges connecting those nodes. There are explicitly drawn undirected edges connecting tones with vowels (e.g., L is connected to [u]). A tone is associated with a certain vowel via the property of being directly connected to that vowel. In addition, there are implicit, undrawn, directed edges between each of the phonemes, since they have a precedence relationship with each other (e.g., [b] precedes [u] in this word).

(10) *Odden (2005:309, example 25) with implicit directed edges added*



With the implicit directed edges of the string at the bottom of the graph added, it is quite easy to see why the above graph is not a tree. Trees are required to be acyclic, and the graph above contains several cycles; see below for one such cycle highlighted in red.

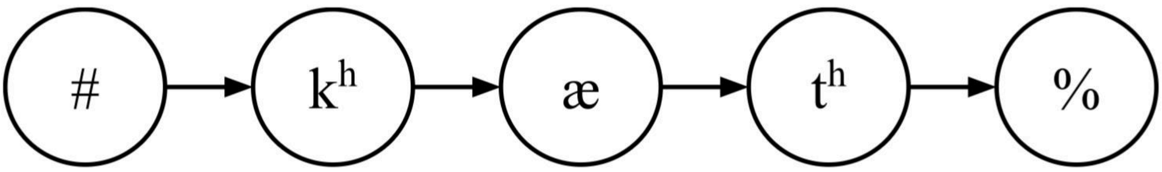
(11) *An implicit cycle in Odden (2005:309, example 25)*



The *Events, Features, and Precedence* (EFP) framework used in this dissertation builds directly off of autosegmental phonology, and like autosegmental phonology, makes use of non-tree graphs to represent phonological phenomena. In the case of EFP, we will be using directed graphs which do allow for cycles. Cycles in EFP are distinct from the ones identified in (11); in EFP-type representations, cycles are typically used to represent reduplication (more detail in the next section and in Chapter 5, see also Raimy 2000 and Papillon 2020).

A very simple EFP graph example of the English word [k^hæt^h] is shown in (12) below. As in (11) above, the direction of an edge in this graph signifies the precedence relationship (either 'precedes' or 'follows') between the nodes it connects.

(12) *The English word [k^hæt^h] using EFP graph*



precedes k^h
k^h precedes æ
æ precedes t^h
t^h precedes %
output: [k^hæt^h]

Since directed graphs in this case are representing temporal phenomena (speech), nodes may also be referred to as events. Events are represented using IPA characters in a straightforward manner. The purpose of # and % may be less intuitive. The # node marks the beginning of the graph and % marks the end. These are sometimes referred to as a start node and end node, respectively. From a network flow perspective, the # node is called the source and the % node is called the sink. In this work, I will be referring to these nodes as the source and the sink, consistent with the idea that EFP graphs are best conceptualized as flow networks (see 2.1.3 for more details).

These nodes are important for identifying where a word begins and ends: when a word is pronounced, flow through the graph must begin at the source and end at the sink. The source and sink nodes are also important for affixation, as they allow us to represent the manner in which suffixes and prefixes attach to a word. Prefixes must follow a source, and suffixes must precede a sink.

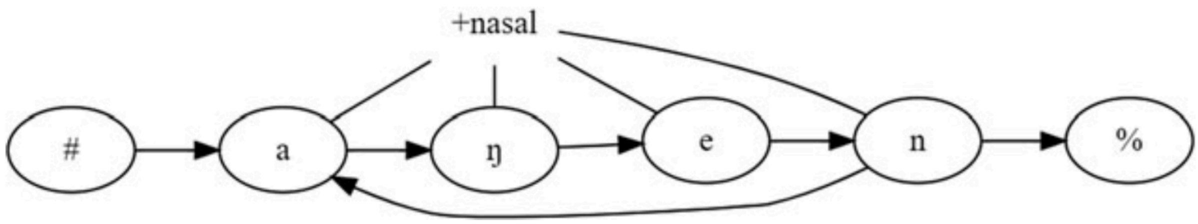
The example in (12) is maximally simple, so it does not appear any different than a typical string. In fact, this example *is* a string, because a string is a very constrained type of directed graph. Strings, as we use them in phonology, are connected, directed tree graphs with no branching- every event is preceded by exactly one event (excluding the source #, which is preceded by nothing) and followed by exactly one event (excluding the sink %, which is followed by nothing). Thus, strings allow only one path from the source to the sink.

If we needed to encode autosegmental information, such as vowel harmony, our directed graph of [k^hæt^h] would no longer be a string. The graph would contain additional edges which are parallel to the de-autosegmentalized phonemes, above and/or below # → k^h → æ → t^h → %. These parallel edges are referred to as parallel autosegmental streams. We will talk a bit more about that in this section, and also in Chapters 3 and 6. In addition, our graph may need to encode reduplication, in which case a cycle would be added to the graph, which would also mean it is not a string. We will discuss that in this section as well, and in Chapter 5.

2.1.2 Previous works using EFP-type phonological representations

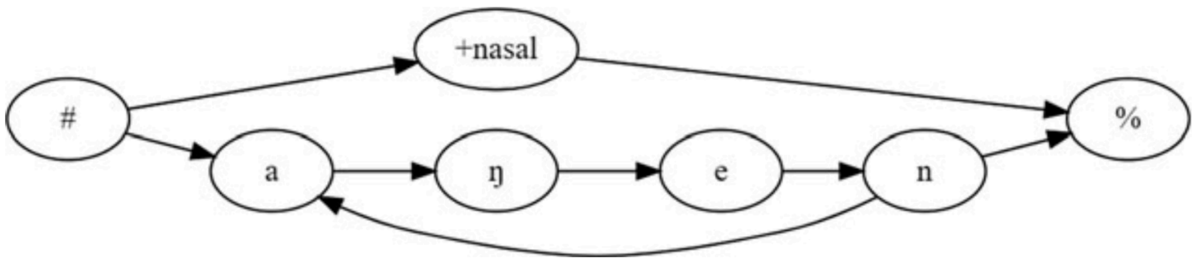
Raimy (2000) introduced the use of directed graphs to describe phonological and morphological phenomena, called multiprecendence. This work contends that lexical items, which are typically represented as strings, are better represented as directed graphs made up of precedence relationships, which allow for cycles and branching. In particular, Raimy (2000) focuses on how directed graph structure lends itself very well to representing reduplication. Below is an example from Malay of the reduplicated word [ãŋẽŋ-ãŋẽŋ], which also contains an autosegmental [+nasal] feature. The autosegmental information spreads via fairly conventional autosegmental feature spreading, associating with different timing nodes (see Goldsmith 1976). The reduplication is applied via a link from the final event, which contains the phoneme [n], to the first event, which contains the phoneme [a].

(13) *Graph of Malay word [ãŋẽŋ-ãŋẽŋ] (from Raimy 2000:18)*



Papillon (2020) further built on Raimy (2000)'s multiprecedence to place autosegmental information on parallel edges (called streams) of the graph. Papillon labeled his updated framework Precedence Relation Oriented Phonology (PROP), and demonstrated how it can represent a diverse array of linguistic phenomena across various languages. See (14) below for a reinterpretation of the graph from Raimy (2000) in (13) above. Note that the autosegmental [+nasal] feature is now located on a parallel stream from the deautosegmentalized information. We will return to representing autosegmental information and reduplication in Turkish in Chapters 3 and 5, respectively.

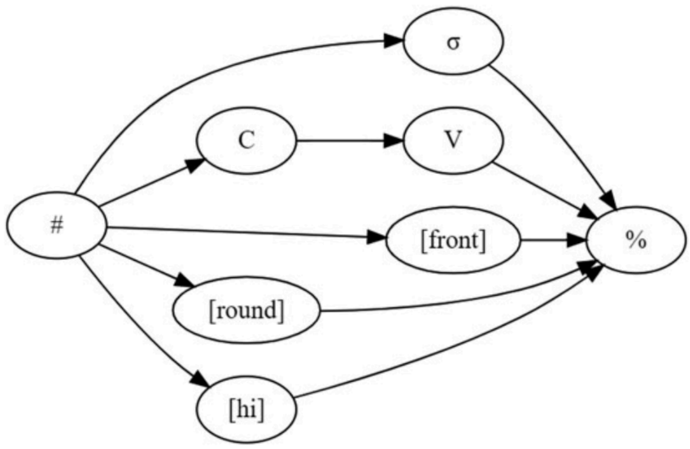
(14) *Graph of Malay word [ãŋẽŋ-ãŋẽŋ] using PROP (from Idsardi 2022:676)*



The *Events, Features, & Precedence* model (EFP; Idsardi 2022) further builds on PROP by laying out a general set of tools for creating phonological representations. Under this framework, three components are necessary to represent phonology: events, features, and precedence. Events are abstract points in time. These events typically contain groups of phonological features. Finally, events are connected via precedence relationships (e.g., event A < event B if A precedes B). Crucially, precedence relationships do not imply a strict timing (e.g., A < B implies only that A precedes B; B need not start immediately after A stops). These three components can be combined into a directed graph representation in a similar manner to Papillon (2020). The EFP model posits that these graph structures (which capture events, features, and precedence) are capable of representing *all* phonological phenomena.

One novel contribution of Idsardi (2022) is that he demonstrates how EFP can be used to represent syllable structure. Discussion of syllable structure is notably brief and inconclusive in Papillon (2020), and Papillon offers this up as an area for future research. Idsardi (2022) argues that like autosegmental information covered by Papillon (e.g., vowel harmony and lexical tone), syllable structure is encoded as discrete events that are on parallel streams to de-autosegmentalized information. Idsardi (2022) elucidates syllable structure in EFP through an example from child development. Baier et al. (2007) found that despite being able to distinguish [i] from [u], and [wi] from [i], two-month-old infants could not distinguish between [ju] and [wi] syllables (see Idsardi 2022 for more detail on this puzzling finding not covered here). Idsardi's solution was to create an EFP graph structure (shown in 15 below) in which [front], [round], and [high] are not sequenced with respect to each other, but rather all occur on parallel streams.

(15) *EFP representation of ambiguous [ju] ~ [wi] syllable (from Idsardi 2022:678)*



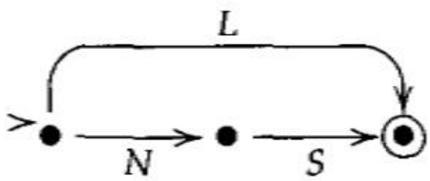
As a result, infants who have this representation cannot distinguish between a [high] syllable with that is [round] followed by [front] ([wi]) and one that is [front] followed by [round] ([ju]). More importantly to this discussion, the graph below also encodes syllable structure. The syllable and its components (onset and nucleus) are each stored in parallel streams. Syllable structure in EFP representations will be revisited in Chapter 4, where a brief proposal for encoding Turkish syllable structure will be laid out.

2.1.3 Further clarification on representation structure

For those readers who have some experience with directed graphs: note that these EFP directed graphs are *not* finite state automata (FSAs). The events are not traversed one by one, and the different streams are not mutually exclusive. For example, if the above graph in (14) was an FSA, an accepted string for this lexical item would be simply [nasal], which is incorrect. It would also be acceptable to infinitely reduplicate [aŋẽn], which is not allowed in real human language.

This representational difference is well typified by an example from Bird & Ellison (1994). Bird & Ellison (1994) argue that FSAs should be used to represent OT constraints and rules. They briefly bring up the possibility of using arc-labeled devices to represent tones which span over multiple segments, as shown in (16):

(16) *Arc-labeled device from Bird & Ellison (1994:62)*



The graph in (16) shows a labial autosegment L which scopes over both a nasal N and a stop S. The idea being that this captures a single labial autosegment applying rounding to both the nasal and the stop. This type of representation is extremely similar to what is used to represent autosegments in EFP, where autosegments are stored on edges or streams that are parallel to de-autosegmentalized information (see the [+nasal] autosegment in the graph in 10, for example). Bird & Ellison (1994) write the following on this type of representation:

> 'However, this representation is flawed: the semantics assigned to coterminous paths contradicts the standard interpretation of FSAs. The automaton pictured above would normally be interpreted as *either* a nasal followed by a stop *or* by a single or by a

> single labial articulation. **Crucially, it could not be interpreted as necessarily *both* a nasal followed by a stop *and* a labial articulation.**' (Bird & Ellison 1994:62, bolding added)

It is entirely true that interpreting (16) as an FSA leads to the incorrect output, and Bird & Ellison were correct to reject it, as their goal was to use FSAs to represent phonology. However, EFP uses this style of representation in its directed graphs, because EFP is not a framework which uses FSAs.

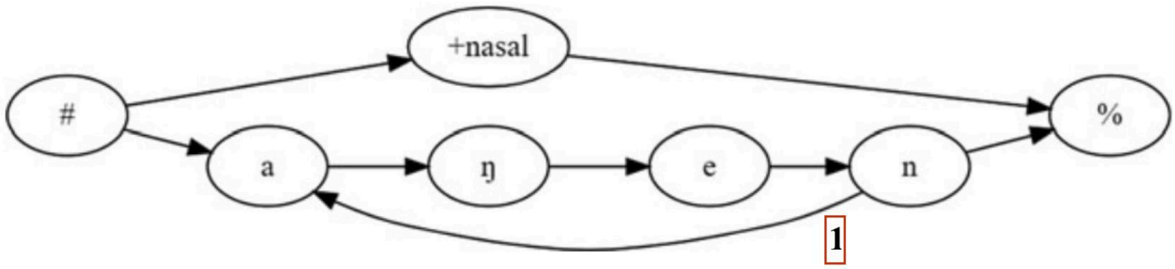
Rather, EFP graphs should be thought of as *flow networks*, whose edges can have a *capacity*. Flow networks are a different kind of directed graph which is interpreted very differently from FSAs. Edges in flow network graphs are often analogized to a set of pipes, through which a certain maximum amount of water can pass (the capacity). These flow networks have a range of uses in mathematics, computer science, and beyond, including creating matchings between sets, optimizing data movement in computer networks, and designing electrical circuits. When using flow networks to represent language, we add one additional constraint which is not typically used in general flow networks: where possible, every edge must have some flow pushed through it (i.e., there is a pressure against unused edges). That is, if there is one possible flow which uses every edge and another that does not, the flow which uses every edge is selected. If there are two possible flows which do not use every edge, then either flow may be used; this represents a choice on behalf of the speaker[3].

While we do not usually explicitly label the capacities of the edges in EFP graphs, since it is fairly intuitive in language examples, capacities can be added without issue, as in (17) below, which has a capacity of 1 added to the reduplicating edge (shown in a red box).

---

[3] The explanation here is quite brief because providing a more depth analysis of how exactly surface representations are computed off of these graphs (in the world's languages in general) goes far beyond the scope of this dissertation, which focuses specifically on Turkish.

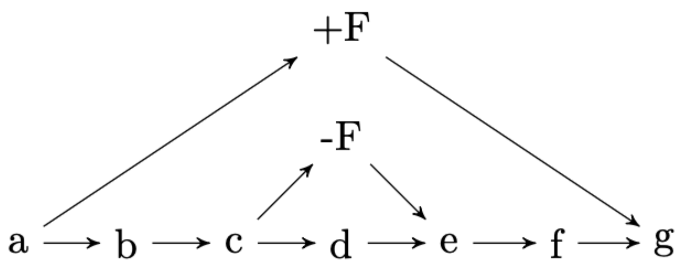
(17) *Graph of Malay word [ãŋẽŋ-ãŋẽŋ] (from Idsardi 2022:676) with capacities added*



When flow occurs through the graph above, flow can be pushed from the source (#) to the sink (%) via the top path (#-[nasal]-%). Further, flow can be pushed via the bottom path (from # to [n]). However, there is a choice at the [n] node. First, there is a more direct path from source to sink (#-a-ŋ-e-n-%). However, recall that the constraint a flow in which every edge is used is preferred over one which has some unused edges. This path (#-a-ŋ-e-n-%) leaves the reduplicating link (n-a edge at the bottom) unused. Hence, we must use the second choice, which is to take a more circuitous path using the reduplicating edge (#-a-ŋ-e-n-a-ŋ-e-n-%). Choosing this path at the [n] node uses the one unit of capacity on the n-a reduplicating edge. The capacity on the reduplicating edge thus fairly straightforwardly avoids the issue of a cycle in the graph allowing for infinite reduplication.

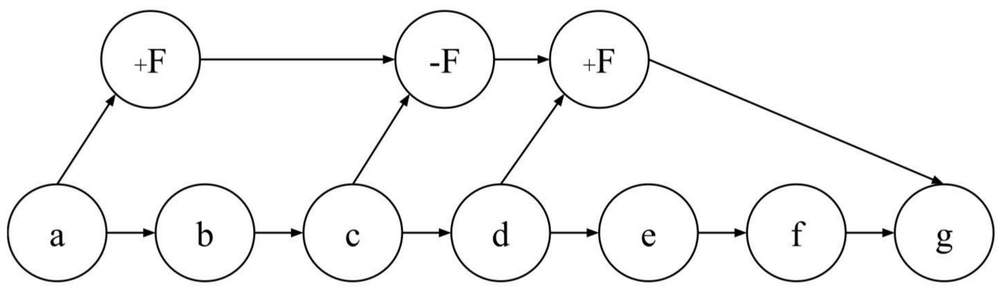
Finally, I would like to briefly discuss the ways that the application of directed graphs in this dissertation will differ from that of Papillon (2020). Firstly, when representing harmony systems, Papillon nests autosegmental information of the same type, with some autosegmental streams for the same feature being nestled under the scope of others, as shown in (18) below.

(18) *Nested autosegmental streams (Papillon 2020:29, fig 23)*



In the above example, the segments {a, b, c, e, f, g} would have feature F, while segment {d} would not. For representations of Turkish vowel harmony, I will eschew this nested structure in favor of using precedence relationships between autosegmental features in a single stream, as shown below in figure (19). I will lay out the reasons I make this change in Chapter 3.

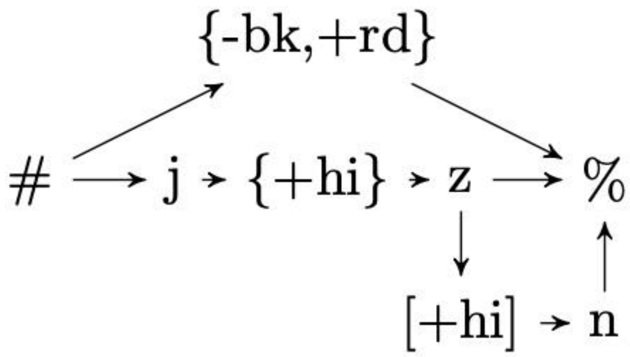
(19) *Non-nested autosegmental information*



The scope of Papillon (2020), in terms of languages, is far broader than that of this dissertation, so it is entirely possible that nested autosegmental streams as in 19 above are necessary for representations in other languages (such as the other vowel harmony systems which Papillon discusses). However, that is beyond the scope of this dissertation to discuss.

Another representational difference between this work and that of Papillon (2020) is in regard to suffixation. In Papillon (2020), when suffixes are added to a base, the link between the final phoneme of the base and the sink is conserved such that both the base and the suffix immediately precede the sink. An example is shown below in (20).

(20) *yüzün [jyz-yn] 'your face' from Papillon (2020:27)*



In the structure given in (20) both [z] (the final consonant of the base) and [n] (the final consonant of the second person single possessive suffix) have an edge which connects them to the sink (%). In this work, suffixes absorb the sink from the base, and add a new sink which comes with their representation. This gives them the appearance of 'slotting in' between the base and the sink. Therefore, only the edge from the final suffix phoneme [n] would be present in the presentation above. This change will be further explained in Chapter 3.

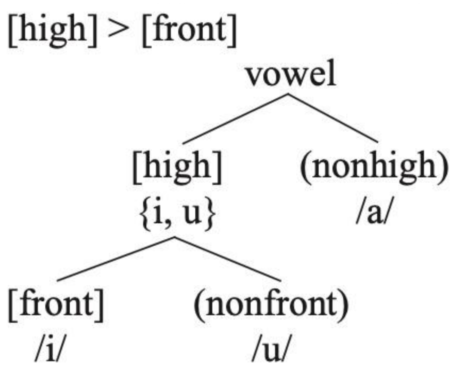
The final key update I will make to Papillon's PROP framework is in regard to the type of features used. Papillon (2020) uses binary features, which are generally fully specified. I will make use of privative features and underspecification, which I explain in more detail in the following section.

2.2 Feature system for this work

In this dissertation, features from Avery & Idsardi (2001) will be used, alongside ideas from the *Modified Contrastive Specification* framework for defining phonological contrastiveness and hierarchy among features (MCS; see Dresher 2008).

MCS is the specific kind of phonological underspecification which is produced by using the Successive Division Algorithm (SDA; see Dresher 2008). The algorithm begins with no contrasts- just a single phoneme. The SDA then successively splits the phonemes into separate categories via one contrastive feature at a time. Anytime a phoneme is already in a category by itself, it ignores the features used to divide up the groups of remaining sounds, marking them as redundant. This process continues until all phonemes are the only member of a group. MCS can also capture hierarchy among contrastive features, where features used to divide phones earlier occur higher up in the hierarchy. As an example, Purnell et al. (2019:e448-e449) show different possible contrastive hierarchies for a language with the three vowels /i ɑ u/. One possibility is that the SDA divides the phonemes first by height, and then by frontness. This results in the contrastive hierarchy shown below in (21).

(21) *Possible contrastive hierarchy for /i a u/ from Purnell et al. (2019:e448, example 2b)*



Note that in the contrastive hierarchy defined above, rounding is not present. This is because height and frontness alone are sufficient to separate all phonemes. In this contrastive hierarchy, rounding is a redundant feature. Whether this ranking, or another (dividing by roundness and then frontness, for example) is correct for a given language with this vowel inventory is dependent on the phonological facts of that language. For more detailed information on MCS and the SDA, see Dresher (2008) and Purnell et al. (2019).

While contrastive hierarchies for features in Turkish are not given in this work, some ideas from MCS do inform the phonology framework employed herein. For example, the idea that phonology is maximally underspecified and only includes features with phonological (rather than purely phonetic) consequences is central. In addition, the idea that there can be features present in phonetics which are redundant is important as well. This stance will be elaborated more below.

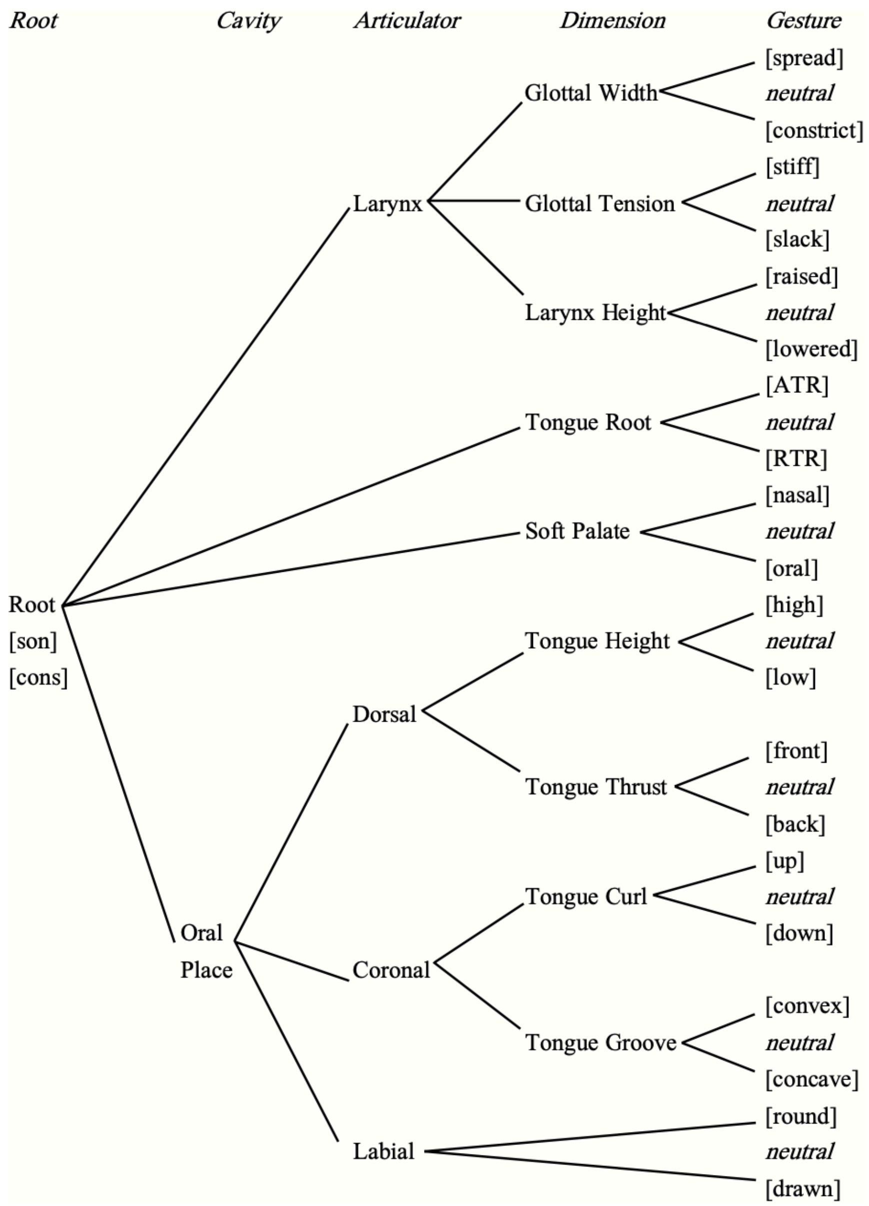
The main system for phonological features in this work is laid out byAvery & Idsardi (2001). This framework incorporates the fact that hierarchically ordered articulators produce speech using distinct muscle groups–known as dimensions–which move in antagonistic ways (see Purnell, Raimy, & Salmons 2019; and Avery & Idsardi 2001 for more detail than is provided here). Avery & Idsardi (2001) organize these dimensions into a feature geometry[4], see the full geometry in (22) below.

For example, in the larynx, the glottis can be actively spread or constricted during a particular speech sound via rotation of the cricoarytenoid joint (Li & Raimy 2017). Additionally, the width of the glottis may remain in neutral position when the cricoarytenoid joint is not actively moved. Note that the glottis must always be in exactly one of three states in terms of its width at any given point in time: spread, constricted, or neutral. These three states are mutually exclusive. For this example, the relevant possible gestures would be [spread] and [constrict], or the absence of a gesture– *neutral*. The dimension would be Glottal Width (GW), under the Larynx articulator.

[4] See Uffman (2011) for a summary and discussion of feature geometries.

Phonological representations always consist of dimensions, which are translated into a gesture during speech via a completion rule. Completion rules can vary cross-linguistically. In English, for example, the GW dimension is used in stops, which surfaces as [p^h t^h k^h] in onset position. As is evident from the transcription, this dimension is completed to a [spread] gesture in English onsets, leading to aspiration. In other languages, the GW dimension is completed to [constrict], which leads to ejective consonants.

(22) *Full feature geometry as shown in (Purnell, Raimy, & Salmons 2019:e465), originally from Avery & Idsardi (2001:66)*



Additionally, note that this framework is privative. For a feature with contrast in a given language, only one of the two gestures associated with a dimension will be used. For example, English has two stop series: <p t k> /p^h t^h k^h/ and <b d g> /p t k/. The former is represented phonologically with a GW dimension (which completes to [spread] in onsets, as mentioned above). The latter has no dimension specification, and receives only a bare articulator and dimension node, LAR^{GW}. This is known as a superordinate marking (Purnell, Raimy, and Salmons 2019). LAR^{GW} stops have the larynx in a neutral position, as in a plain unaspirated stop.

Lastly, as mentioned earlier, the feature system used in this dissertation is very underspecified in the phonology. I follow the idea from Avery & Idsardi (2001:41) that 'representational economy forces the phonology to be non-redundant in the sense that phonological representations are minimally specified' and that 'unlike the phonological representations, the phonetic representations are tremendously over-specified, containing information far beyond what is necessary for the simple maintenance of contrast.'

For example, as mentioned earlier, English uses GW to maintain stop contrasts, but speakers of certain dialects (like Southern American English) pronounce the plain series <b d g> /p t k/ with prevoicing, as [b d g] (see for example, Walker 2020). This is because these speakers enhance their plain stops with prevoicing, using glottal tension (GT). Glottal tension, while used in some Southern American English speech, is not present as a dimension in the phonology. It is present only in the phonetics to enhance the phonological aspiration contrast. Dimensions are translated into gestures during speech using completion rules. The completion rule for the English GW stops in onsets is shown below:

(23) *Completion rule for Glottal Width in onsets*

Glottal Width (GW) → [spread] (spread vocal folds, facilitating aspiration)

As mentioned above, superordinates can also be enhanced in the phonetics. Below is an enhancement rule for plain unaspirated stops in Southern American English.

(24) *Enhancement rule for plain stops in Southern American English*

Larynx^{GT} (LAR^{GT}) → [slack] (slack vocal folds, facilitating voicing)

Unlike the completion rule in (23), the enhancement rule in (24) is not present in the phonology- only in the phonetics. As mentioned before, in this framework, phonology is minimally specified, while phonetics is over-specified, as is the case with the Southern American English GW contrast in stops.

2.3 Exceptions to Turkish Vowel Harmony

This section of the literature review will cover necessary background information on Turkish vowel harmony (TVH), as well as so-called 'exceptions' to TVH. First, general background on Turkish vowels and TVH will be discussed for any readers not already familiar, followed by a review of relevant literature on TVH and its exceptions.

2.3.1 Background on Turkish vowels

Turkish has eight vowels with three contrastive features: height, rounding, and frontness. The vowel inventory is quite symmetric: there are four high vowels, four rounded vowels, and four front vowels. All eight vowels are organized by contrastive features into the chart below:

(25) *Turkish vowels (from Göksel & Kerslake, 2005:9)*

	Front		Back	
	Rounded	Unrounded	Rounded	Unrounded
High	*ü* /y/	/i/	/u/	*ı* /ɯ/
Mid and Low	*ö* /œ/	/e/	/o/	*a* /ɑ/

Vowels can be either high or non-high. There are 'mid' vowels (e.g., /o/ and /e/) but the mid/low difference is not contrastive, as exemplified by the free variation between /e/~/æ/ before sonorants and the alveolar tap (Göksel & Kerslake 2006:10). For frontness, the vowels can be separated into front and back. Turkish has no central vowels.

2.3.2 Background on Turkish vowel harmony

TVH operates on frontness and rounding. Height is lexical and not subject to harmony. All vowels are subject to frontness harmony, while rounding harmony is active only on high vowels. TVH spreads features from left to right across a word. Therefore, in roots, all vowels should match the preceding vowel in frontness and, if the vowel is high, rounding. The following shows several examples of multisyllabic roots with vowels that conform to TVH:

(26) *Harmonic roots*

	Word	*IPA*	*Translation*
a.	övünç	[œvyntʃ]	'pride'
b.	hediye	[hedije]	'cat'
c.	okul	[okul]	'school'
d.	sarımsak	[sɑrɯmsɑk]	'garlic'

TVH also applies across morpheme boundaries, spreading features from root vowels to suffix vowels[5]. Since high and non-high vowels differ in whether they are subject to rounding harmony, high harmonizing suffix vowels will be written as a capital *I* and low harmonizing suffix vowels will be written as a capital *A* in the bare form. Below is an example of a low harmonizing suffix vowel with the plural suffix -lAr:

(27) *Frontness harmony on a low suffix vowel*

	Word	*IPA*	*Translation*
a.	kedi-ler	[kedi-ler]	‘cats’
b.	okul-lar	[okul-lɑr]	‘schools’

Since non-high vowels are not subject to rounding harmony, rounding in the preceding root vowel does not affect the plural suffix vowel (compare: *okullar* [okul-lɑɾ], **okullor* *[okul-loɾ]).

Rounding harmony does apply to high vowels, so high vowel suffixes harmonize for both rounding and front/backness. This can be seen in the examples with the second person singular possessive suffix /-In/ below:

(28) *Rounding harmony in a high suffix vowel*

	Word	*IPA*	*Translation*
a.	övünç-ün	[œvyntʃyn]	‘your (sg) pride’
b.	okul-un	[okulun]	‘your (sg) school’
c.	bank-ın	[bɑnkɯn]	‘your (sg) bench’
d.	köpeğ-in	[kœpejin]	‘your (sg) dog’

[5] Almost all affixes in Turkish are suffixes. The language has very few loan prefixes which do not appear to undergo vowel harmony with the base. For now, I will put these to the side, and we will return to them later in subsection 3.4.1.

2.3.3 TVH is opaque: rounding opacity

Turkish vowel harmony is opaque, meaning that a harmonizing vowel will always harmonize with the immediately preceding vowel, as opposed to a transparent system such as Finnish, in which harmony may 'skip' certain vowels (van der Hulst & van de Weijer 1995).

One notable natural consequence of TVH being opaque is rounding opacity, in which non-high vowels block rounding harmony. Because rounding harmony is not active on non-high vowels, non-high vowels can block the spread of rounding to high vowels. Below is a classic example, also given in Papillon (2020):

(29) *Rounding Opacity*

	Word	*IPA*	*Translation*
a.	yüz	[jyz]	'face'
b.	yüzün	[jyz-**yn**]	'your face'
c.	yüzler	[jyz-leɾ]	'faces'
d.	yüzlerin	[jyz-leɾ-**in]**	'your faces'

The crucial alternation is in the second person singular possessive suffix (bolded for clarity). In example (29b), the suffix (consisting of a single high vowel vowel followed by an /n/) harmonizes with both the backness and rounding of the preceding vowel *ü* [y]. In example d, the rounding harmony is blocked by the plural suffix *-lAr* /-lAɾ/. Because the plural suffix contains a non-high vowel, it does not undergo rounding harmony and the vowels following it will not undergo rounding either. This blocking process is known as rounding opacity, and it is quite common in Turkish. Many suffixes have low harmonizing vowels which block rounding harmony in the same manner demonstrated for *-lAr* /-lAɾ/

above, including *-(y)A* ‘to’, *-mA* ‘do not’, *-ArAk* gerund of state, *-DA* ‘in, at, on’, *-DAn* ‘from’, and others[6].

2.3.4 Background on disharmony / ‘exceptions’

There are many Turkish words which do not conform to the generalizations given about vowel harmony above. This section contains information about these words.

The first type of ‘exception’ (the most common) is an immutable vowel. In this case, a vowel within a morpheme (root or suffix) will never change, causing the morpheme to not conform to TVH. This type of exception is very common. A corpus study found that the first two vowels of a root are mismatched for frontness about one third of the time (Kabak et al. 2008).

Below, there is a table with all possible unique combinations of two sequential vowels in Turkish. Where a sequence is attested, a word is added to the relevant cell. The entries are color-coded. Green cells denote a combination which conforms to vowel harmony. Yellow cells indicate that the combination does not conform to frontness harmony. Orange cells denote that the combination is not harmonic for rounding (note that only high vowels harmonize for rounding). Finally, red indicates that a combination is not harmonic for frontness or rounding. For an alphabetic list and translations of the words to English, see Appendix 1.

[6] The capital letter D on some of these suffixes denotes that the consonant surfaces as either [d] or [t], assimilating to the voicing of the preceding phone.

(30) *Table of Disharmonic Turkish Roots*

V2 →

V1 ↓	a /ɑ/	ı /ɯ/	o	u	i	e	ü /y/	ö /œ/
a /ɑ/	para /pɑɾɑ/	sarı /sɑɾ/	tablo /tɑblo/	armut /ɑɾmut/	hani /hɑni/	anne /ɑn:e/	tesa**dü**f /tes**ɑ**d**y**f/	asan**sö**r /ɑs**ɑ**ns**œ**ɾ/
ı /ɯ/	sıra /sɯɾɑ/	ışık /ɯʃɯk/	hırbo /hɯɾbo/	haz**ıru**n /hɑz**ɯ**ɾ**u**n/	**ıtri**yat /ɯtrijɑt/	kırlent /kɯɾlent/	**none?**	**none?**
o	oda /odɑ/	rodstır /ɾodstɯɾ/	horoz /hoɾoz/	okul /okul/	dep**ozi**t /dep**ozi**t/	poşet /poʃet/	modül /modyl/	şoför /ʃofœɾ/
u	kupa /kupɑ/	muzır /muzɯɾ/	soru /soɾu/	duygu /dujgu/	kuzin /kuzin/	kubbe /kub:e/	gusül /gusyl/	dublör /dublœɾ/
i	kitap /kitɑp/	printır /pɾintɯɾ/	kilo /kilo/	**ipu**cu /**ipu**dʒu/	iki /iki/	inek /inek/	virüs /viɾys/	fritœz /fɾitœz/
e	elma /elmɑ/	menkıbe /menk**ɯ**be/	sezon /sezon/	cesur /dʒesuɾ/	kedi /kedi/	yeter /jeteɾ/	menü /meny/	profesör /pɾofesœɾ/
ü /y/	m**üca**dele /mydʒɑdele/	**none?**	külot /kylot/	nüfus /nyfus/	ümit /ymit/	mücver /mydʒver/	müdür /mydyɾ/	tümör /tymœɾ/
ö /œ/	rötar /ɾœtɑɾ/	**none?**	nöron /nœɾon/	rötuş /ɾœtuʃ/	rölik /ɾœlik/	böcek /bœdʒek/	kötü /kœty/	r**ölö**ve /ɾ**œlœ**ve/

Almost all of the cells of the table are full, meaning that almost every possible combination of vowels is attested in Turkish. There are some gaps in the table which will be discussed later on. For now, let it suffice that there are many examples of disharmony in roots.

In addition to disharmonic roots, there are a number of suffixes with invariable vowels that do not harmonize. For example, the ability suffix *-(y)Abil* /-(j)Abil/ has one harmonizing vowel and one immutable /i/ vowel. Another common example is the present progressive suffix *-Iyor* /-Ijoɾ/, which has a high harmonizing vowel and an immutable /o/

vowel. Examples of both are shown below (with immutable, non-harmonizing suffix vowels bolded).

(31) *Immutable suffix vowels blocking harmony*

	Word	*IPA*	*Translation*
a.	yürü-yeb**i**l-ir	[jyry-jeb**i**l-iɾ]	'he/she/it can walk'
b.	at-ab**i**l-ir	[ɑt-ɑb**i**l-iɾ]	'he/she/it can throw'
c.	yür-üy**o**r	[jyr-yj**o**ɾ]	'he/she/it is walking'
d.	at-ıy**o**r	[ɑt-ɯj**o**ɾ]	'he/she/it is throwing'

The second vowel of the ability suffix *-(y)Abil* /-(j)Abil/ remains unchanged when it is suffixed to a front vowel word (31a) and a back vowel word (31b). Similarly, the second vowel of the progressive suffix *-Iyor* /-Ijoɾ/ remains [o], regardless of being attached to a word with front rounded vowels (31c) or back unrounded vowels (31d).

In addition to immutable vowels, there are a number of roots with back final vowels that exceptionally condition front suffix vowels. In these cases, harmonizing suffix vowels which typically match the backness of the preceding vowel are exceptionally front when affixed to these roots. A number of examples are shown below:

(32) *Roots which take exceptional front vowel suffixes*

	Word	*IPA*	*Translation*
a.	roll**e**r	[rol-l**e**ɾ]	'roles'
b.	petrol**ü**	[petɾol-**y**]	'oil-ACC'
c.	saat**i**	[sɑɑt-**i**]	'clock-ACC'
d.	goll**e**r	[gol-l**e**ɾ]	'goals'
e.	tuval**i**	[tuval-**i**]	'canvas-ACC'
f.	harf**i**	[hɑɾf-**i**]	'letter-ACC'
g.	kalpl**e**r	[kɑlp-l**e**ɾ]	'hearts'
h.	idrak**i**	[idɾɑk-**i**]	'perception-ACC'
i.	dikkat**i**	[dikkɑt-**i**]	'caution-ACC'

This case differs from the immutable vowels in the previous section, because these suffix vowels are not immutable- they typically do conform to TVH (e.g., *kitabı* [kitabɯ] 'book-ACC'). These exceptions are much less numerous than the previous root and suffix exceptions discussed in the previous section in that there are relatively few roots which show this pattern. However, this kind of exception is not at all rare in the sense that there are high-frequency back vowel words which take front vowel suffixes (e.g., *saat* [sɑɑt] 'clock', *alkol* [ɑlkol] 'alcohol', *hayal* [hɑjɑl] 'dream').

In addition to exceptional front suffixes two sources (Clements & Sezer 1982, Lewis 1967) report that a small number of Arabic loanwords with front vowels exceptionally take back suffixes, as shown below:

(33) *Roots which take exceptional back vowel suffixes (Clements & Sezer 1982)*

	Word	*IPA*	*Translation*
a.	sevkı	[sevk-**ɯ**]	'desire-ACC'
b.	fevkı	[fevk-**ɯ**]	'top-ACC'
c.	halikı	[hɑlik-**ɯ**]	'creator-ACC'
d.	tasdikı	[tɑsdik-**ɯ**]	'confirmation-ACC'
e.	utaridı	[utɑɾid-**ɯ**]	'Mercury (planet)-ACC'

Clements & Sezer (1982) specify that exceptions of this type are found only in 'some idiolects' and amongst 'mostly older' speakers (pp 242). A follow up study also reported in Clements & Sezer (1982) found that, of five Turkish speakers tested, none provided back vowel suffixes for the roots above (pp 254-255). As such, these forms are likely obsolete, but an analysis will nonetheless be provided in Chapter 3.

2.3.5 Exceptions and opacity

As noted in the earlier discussion of rounding harmony blocking, TVH is opaque. This opacity also applies in the case of 'exceptional' vowels. Like non-high vowels, vowels that do not conform to TVH will block the spreading of features. Vowels which follow the 'exceptional' vowel will conform to the harmony of the 'exceptional' vowel and the spreading of frontness and/or rounding from vowels earlier in the word will be blocked. The set of words below exemplifies this for 'exceptional' suffix vowels:

(34) *Immutable suffix vowels blocking harmony*

	Word	*IPA*	*Translation*
a.	içiyorum	[itʃ-ijoɾ-**um**]	'I am drinking'
b.	içerim	[itʃ-eɾ-**im**]	'I drink'
c.	atıyorum	[ɑt-ɯjoɾ-**um**]	'I am throwing'
d.	atarım	[ɑt-ɑɾ-**ɯm**]	'I throw'

As noted in the preceding section, the second vowel of the progressive suffix *-Iyor* /-Ijoɾ/ does not undergo any kind of harmony with the preceding vowel; it will always surface as [o]. The aorist suffix -Ar and the first person singular subject agreement suffix -Im contain a harmonizing low and high vowel, respectively. The singular subject agreement suffix -Im is bolded; note the different surface forms in the example set.

As an example, for the form in (34a) *içiyorum* [itʃ-ijoɾ-**um**], the immutable vowel [o] blocks the spreading of frontness to the first person subject agreement suffix -Im, causing it to surface as a back vowel. When the aorist suffix -Ar is added instead as in (34b) *içerim* [itʃ-eɾ-**im**], frontness is able to spread to both suffixes.

For those roots which take exceptional front suffixes, opacity applies as well. The exceptionally front suffix will block the spreading of backness from the root, as shown in (35) below.

(35) *Exceptional front suffixes blocking the spread of backness*

a. saatine [sɑːt-in-e] 'to your (sing) clock'

cf., *saatina [sɑːt-in-**a]**

b. gollerim [gol-leɾ-im] 'my goals'

cf., *gollerım [gol-ler-**ɯ**m]

When a back vowel word, such as *saat* [sɑ:t], takes an exceptional front vowel suffix, subsequent harmonizing suffixes will be front. The exceptional front suffix is not transparent and backness from the root will be blocked from spreading to any of the suffixes.

2.3.6 Past accounts of exceptions to TVH

This section will briefly review some of the phonological literature about Turkish vowel harmony. Lightner (1965) is an early example of work which seeks to sort out apparent anomalies in the vowel harmony system of a Turkic language. Lightner (1965) discusses vowel harmony in Classical Mongolian. Classical Mongolian has vowel harmony that applies to frontness; that is, vowels of a given word will be front or back, except for [i], which can co-occur with front or back vowels. Specifically, Lightner (1965) proposed an analysis for a seeming oddity in some Classical Mongolian words: some roots for which all vowels were [i] would take front vowel suffixes, while others would take back vowel suffixes. Lightner concluded that rather than being the result of some phonological rule, entire roots are lexically marked as either front or back, and this marking determines the front harmony on suffix vowels. The phonological representations laid out for TVH in this dissertation are very much in the spirit of Lightner (1965)'s account for Classical Mongolian. Indeed, for the representations of words which conform to TVH, we have marked roots with TT, LAB, etc, autosegmental spans that mark whole roots with their frontness and rounding features- very much in line with Lightner (1965)'s proposed analysis of Classical Mongolian. This line of thinking will continue with the discussion of exceptions later on.

Clements & Sezer's seminal 1982 paper was one of the first to seriously consider the phonology of exceptions to TVH specifically, and remains oft-cited to this day. Clements & Sezer (1982) described TVH using an autosegmental framework (see Goldsmith 1976) and discussed explanations for apparent exceptions.

Clements & Sezer (1982) address disharmonic roots as follows. First, they claim that the vowels /ɑ e i o u/ must be able to occur freely within roots, that vowel harmony does not act on them in this environment. This argument is supported by the fact that, as discussed in subsection 2.3.4, a high number of words with these vowels do not conform to TVH (e.g., *kitap* [kitap] 'book', *mektup* [mektup] 'letter', *anne* [an:e] 'mother', etc). In addition, in the nonce word experiments they conducted, Turkish speakers sometimes prefer disharmonic words with these vowels over words with harmonic vowels (pp 225, 227).

Clements & Sezer (1982) further argue that *ı ü ö* /ɯ y œ/ (with the exception of *ü* /y/ when used with /i/) are always subject to harmony rules within roots, despite the apparent exceptions to TVH involving these vowels. Exceptions involving these vowels are generally less common. In addition, exceptions involving the vowels *ı ü ö* /ɯ y œ/ were very rarely preferred in the previously mentioned nonce word studies, which was taken as confirmation that speakers are sensitive to harmony in roots with these vowels. Finally, Clements & Sezer (1982) report that word with disharmonic *ı ü ö* /ɯ y œ/ are regularized different extents in different dialects of Turkish. Here are some examples from their paper, along with regularized forms:

(36) *Regularization of disharmonic ö ü /œ y/ (from Clements & Sezer 1982:223)*

	Word	*IPA*	*Translation*
a.	ş**o**för ~ ş**ö**för	[ʃ**o**fœɾ ~ ʃ**œ**fœɾ]	'chauffeur'
b.	n**ü**fus ~ n**u**fus	[n**y**fus ~ n**u**fus]	'population'
c.	kom**ü**nist ~ kom**i**nist	[kom**y**nist ~ kom**i**nist]	'communist'
d.	ş**ö**valye ~ ş**o**valye	[ʃ**œ**valje ~ ʃ**o**vɑlje]	'knight'

They contend that because these exceptions are not pronounced with the front rounded vowels some amount of the time, they do not constitute true exceptions to TVH in the same way as words with disharmonic /ɑ e i o u/. They argue that this further supports their claim that *ü ö* /y œ/ are subject to vowel harmony within morphemes, whereas /ɑ e i o u/ are not.

Clements & Sezer (1982) argued that immutable suffix vowels, like that of the present progressive *-Iyor* /-Ijoɾ/, are opaque vowels. These opaque vowels are specified for frontness and rounding in autosegmental tiers underlyingly.

Whether TVH is active within roots, and to what degree, remains a contentious topic. Some have argued, like Clements & Sezer (1982), that only certain vowels are subject to TVH within roots. One such example is Pöchtrager (2010), who argues that only ı /ɯ/ is subject to vowel harmony within roots. However, it has been argued that it is not simple to identify a class of vowels which is not subject to TVH within roots, due to the diverse nature of exceptional root vowels (see Kabak 2011). Additionally, others argue that vowel harmony is no longer active on roots at all, but rather a morphological process that some suffixes undergo (see Polgárdi 1999, Bacanlı et al. 2020).

The analysis seems here to be stuck between a proverbial rock and a hard place. If TVH is not active within roots, there are some facts that become very difficult to explain. There seems to be no way to account for the regularizations noted by Clements & Sezer (1982) above. In addition, a previously mentioned corpus study found that the first two vowels of a word do not match for frontness about 30% of the time (Kabak et al. 2008). Even taking this number at face value, if TVH is not active within roots whatsoever, it is difficult to explain why the first two vowels of a word match for frontness 70% of the time. Given that half of the Turkish vowel inventory is front, and half is back, pure chance would predict a 50% matching rate. Along the same lines, there would be no satisfactory means to explain why some loanwords do become regularized (e.g., *doktor* [doktoɾ] 'doctor' from French *docteur* which has a front rounded vowel in the second syllable, cf *doktör *[doktœɾ]). On the other hand, if one supposes that TVH is active within roots, then one is then tasked with accounting for the large number of root words which do not conform to TVH. It seems that both approaches come with significant drawbacks. This issue is at the heart of the analysis given in this dissertation, and it will be revisited in Chapter 3.

To account for roots which take exceptionally front or back suffixes (see tables 32 & 33), Clements & Sezer (1982) argue that opaque underlying consonants block the spread of frontness harmony from vowels in the root. This seems to be the prevailing understanding of these exceptions (see Kornfilt 1997, Nevins 2010, Özçelik 2024).

In order to understand this view, it is important to understand how consonants are affected by vowel harmony in Turkish. Turkish has front allophones [c ɟ l̪] of the velar consonants [k g ɫ]. Göksel & Kerslake (2006) state that front allophones occur in front vowel environments and velar allophones occur in back vowel environments. Here, the change in consonant place is attributed to the front feature spreading from a front vowel to the velar consonant, causing the consonant to be pronounced as palatal. This phenomenon is sometimes referred to as a type of consonant harmony (Clements & Sezer 1982, Kornfilt 1997).

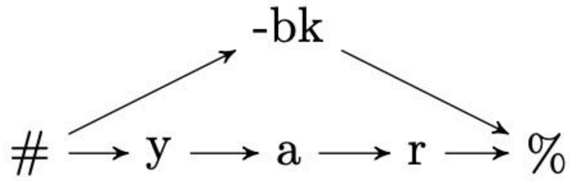
However, not all instances of these palatal consonants are due to allophony. There are some instances wherein a front consonant [c ɟ l̪] occurs next to a back vowel, as in *rüzgâr* [ɾyɟɑɾ] 'wind'. In addition, there are some rare cases where the front and back forms of a velar are in contrastive distribution, e.g., *kar* [kɑr] 'snow' and *kâr* [cɑr] 'profit'. The hat on the back vowel in *kâr* [cɑr] 'profit' marks the preceding <k> as palatal (Turkish term for vowels of this type: şapkalı sesli harfler).

Clements & Sezer (1982) argue that exceptionally front suffixes are conditioned by word-final front consonants that follow back vowels. For instance, *rol* [rol̪] 'role' has a front /l/ which blocks the spreading of the [back] feature from the vowel onto the following suffix, and instead spreads a [-back] feature. Conversely, in the very rare case of exceptional back vowel suffixes, like *şevk* [ʃevk] 'desire', the final consonant is marked as [back], which blocks the spreading of the [-back] feature from the vowel in the base.

While this analysis provides an adequate explanation for those roots that end with /l/ or /k/ which take exceptional suffixes, there are other such roots which end in different consonants. For instance, *saat* [sɑːt]'clock', which takes front vowel suffixes, does not end with /l/ or /k/. To account for this data, Clements & Sezer propose that more consonants, namely /t/ and *r* /ɾ/, have front and back phonemes which also influence vowel harmony. Such consonants are certainly never contrastive and are not attested in grammars of Turkish (see Lewis 1967, Underhill 1976, Kornfilt 1997, Göksel & Kerslake 2006). However, there has been at least one attempt to confirm whether these front and back allophones of /t/, /d/, and *r* /ɾ/ might be produced by speakers which found some potential evidence for their existence based on F2 production data (Canalis & Dikmen 2020). This idea will be further discussed in Chapter 3.

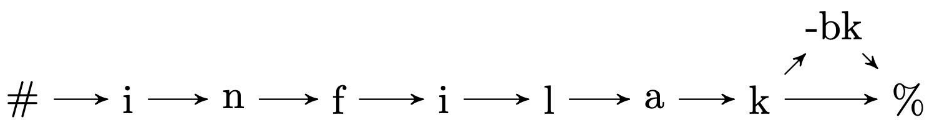
Papillon (2020) accounts for this kind of 'exception' using a parallel autosegmental stream which stores a feature that is opposite to that of the vowel (e.g., [-back] on a root with a back vowel). The graph in (37) below shows a representation of a back vowel root which takes a front vowel suffix from Papillon (2020). In Papillon's account, the root vowel is fully specified, and is not affected by this parallel stream. However, when an underspecified suffix vowel is added onto the root, it falls under the scope of this autosegmental stream. This results in a suffix vowel which is disharmonic with the root vowel, with respect to frontness.

(37a) *Graph of front suffix word yâr [ja:ɾ][7] 'lover' from Papillon (2020:31)*



There is an additional, alternative analysis given by Papillon (2020) for this kind of exception. The alternative analysis is to have a [-back] feature which has scope following the final sound of the root. This means that when the root is suffixed with harmonizing suffixes, they fall under the scope of the [-back] feature, so they surface as [-back] even though the final vowel of the root is [+back]. This second analysis given by Papillon (2020) is very similar to the one proposed in Chapter 3. See (37b) below for an example from Papillon (2020):

(37b) *Graph of front suffix word infilak [infilak] 'explosion' from Papillon (2020:41)*



Another source which has shaped this account is Nevins' 2010 book on harmony systems. Nevins (2010) lays out a search and copy model of vowel harmony in which harmonizing segments seek out harmony information from the closest segment which contains the relevant feature(s). In Turkish, this means that harmonizing vowels seek out

[7] The circumflex in Turkish orthography can be used for multiple purposes. Here the circumflex on the vowel does *not* denote a palatal preceding consonant, as it did for examples *kâr* [cɑr] 'profit' and *rüzgâr* [ɾyɟɑɾ] 'wind'. In this case, the circumflex denotes a long vowel.

frontness and/or rounding from the nearest segment which contains these feature(s). According to Nevins, this segment is typically another vowel, but can be a consonant in some situations. This idea that Turkish vowel segments are somehow ‘seeking’ frontness and rounding specification will be an important part of the account outlined in this dissertation.

In addition, Optimality Theory (OT) is another framework which has been applied to vowel harmony, and to exceptions to TVH in particular. OT will not be used in this dissertation, but since this is an important kind of approach, it will be briefly summarized here. Kabak & Weber (2013) characterize OT approaches to harmony & disharmony as falling into two broad categories: directionless constraints of the agree family and directional constraints of the spread type. These approaches use some kind of high ranking faithfulness constraint to account for exceptions. This faithfulness constraint is ranked higher than harmony constraints, which blocks harmony. As an example of the former OT approach, Polgárdi (1999), who argues that exceptions to TVH occur because TVH is only active across morpheme boundaries, posits a derived environment faithfulness constraint which assigns violations when TVH is applied within morphemes. This constraint outranks the harmony constraint, thereby blocking harmony from applying within morphemes.

As an example of the second approach, Kirchner (1993) uses directional constraints. He posits an align constraint which spreads the relevant features for TVH bidirectionally across a morpheme. Kirchner then deals with disharmony via a *Spread faithfulness constraint which is ranked above the align constraint. Single features may be pre-linked to multiple vowels with a root, but this is a lexical specification, and not indicative of vowel harmony, and does not violate the *Spread faithfulness constraint. This approach represents Kirchner’s position that vowel harmony is no longer active within roots.

2.3.7 TVH conclusion

This subsection contained a wide overview of issues related to vowel harmony in Turkish. First, general facts about vowel harmony in Turkish were provided, and rounding opacity was covered. Then, examples of two different kinds of 'exceptions' to TVH were explained, along with a discussion of some past approaches to accounting for them. Vowel harmony will be revisited in Chapter 3, where a full account which covers all data, including 'exceptions' will be laid out.

2.4 Vowel alternation in coda clusters

This section covers a ∅ ~ high vowel alternation which occurs in some Turkish words. The main analysis of this data is that the alternating vowel is an epenthetic vowel which is added to repair an illicit coda cluster. There is another approach to the data which is that the alternating vowel is present in the underlying representation, and that this alternation is a result of morpho-lexical syncope. Background information about the alternation, as well as a discussion of each of the past approaches will be given below.

2.4.1 Background on vowel alternation in some codas

There is a kind of potential epenthetic vowel in Turkish which occurs at the ends of root words. There is a ∅ ~ high vowel alternation in word-final consonant clusters of a limited set two-syllable words (in Turkish, this is labelled as one form of *ünlü düşmesi* or 'vowel drop'). Interestingly, many of the words which have this alternation are body parts (Erdal 2010, Özçelik 2024). The bare form is bisyllabic, with a high vowel in the second syllable. However, when the final consonant is resyllabified into an onset via the addition of a vowel-initial suffix, the vowel in the second syllable of the root is not pronounced. Here are a few examples:

(38) *Examples of ∅ ~ high vowel alternation in coda clusters*

	Bare Form	*Vowel-Initial Suffix*
a.	vak**i**t [vɑ.k**i**t] 'time'	vaktim [vɑk.t-im] 'my time'
b.	boy**u**n [bo.j**u**n] 'neck'	boynum [boj.n-um] 'my neck'
c.	sab**ı**r [sɑ.b**ɯ**ɾ] 'patience'	sabrım [sab.ɾ-ɯm] 'my patience'

The alternating high vowel can be disharmonic, as shown in [vakit] 'time' above, with a front high vowel following a back vowel. However, alternating back vowels are never found after front vowels. It is clear that the presence of the second vowel in the root is related to suffixation, as when a consonant-initial suffix is added, and the final consonant of the root is not re-syllabified into an onset, the second vowel must be pronounced:

(40) *Examples of alternating words with consonant-initial suffixes*

	Bare Form	*Consonant-Initial Suffix*
d.	vak**i**t [vɑ.k**i**t] 'time'	vak**i**tten [vɑ.kit.-ten] 'from time'
e.	boy**u**n [bo.j**u**n] 'neck'	boy**u**nlar [bo.jun.-lɑɾ] 'necks'
f.	sab**ı**r [sɑ.b**ɯ**ɾ] 'patience'	sab**ı**rdan [sa.bɯɾ.-dan] 'from patience'

When a suffix which begins with a non-high vowel is affixed on to a stem, some speakers may still include the second vowel in the root. Göksel and Kerslake (2006:18) point out this interspeaker variation, and include the following examples:

(41): *Variation in inserted vowel from Göksel and Kerslake (2006:18)*

a.	*Bare Root*:	karın [kɑ.ɾɯn] 'belly'
	Suffixed w/ high vowel:	karnım [kɑɾ.n-ɯm] 'my belly'
	Suffixed w/ non-high vowel:	karna [kɑɾ.n-a] OR karına [kɑ.ɾɯ.n-a] 'to the belly'
b.	*Bare Root*:	şehir [ʃe.hiɾ] 'town'
	Suffixed w/ high vowel:	şehrim [ʃeh.ɾ-im] 'my town'
	Suffixed w/ non-high vowel:	şehre [ʃeh.ɾ-e] OR şehire [ʃe.hi.ɾ-e] 'to town'

Lewis (1967:9) also points out the fact that some speakers use the form *şehire* [ʃe.hi.ɾ-e] rather than *şehre* [ʃeh.ɾ-e], and predicts that this root is changing such that its second vowel will always be present. Lewis' prediction has not yet come to pass, but his comment serves as further evidence that this variation exists and has been ongoing for quite some time (at least about 60 years).

2.4.2 Past accounts of vowel alternation in some codas

Clements & Sezer (1982) explain the alternation as a repair strategy for extra-syllabic consonants which cannot be added to an onset or coda due to the phonotactics of Turkish. For example, the underlying representation of the root in (39a) above would be /vakt/, but in the bare form a vowel is added to repair the /kt/ consonant cluster, yielding *vakit* [vakit] in the surface form. They posit a rule that 'inserts an epenthetic high vowel, unassociated with any features on the autosegmental tiers involved in harmony, between two consonants if the second can form neither a syllable onset nor a syllable coda *by the syllable structure rules of Turkish*' (Clements & Sezer 1982:246, italics added).

Essentially, in this account, vowel insertion occurs to break up illicit coda clusters at the ends of two-syllable words. This viewpoint seems to be the prevailing one; it is echoed in Kornfilt (1986), Kabak (2011), Özçelik (2024), and is the account given in some grammars of Turkish (see Lewis 1967:19, Ketrez 2012:17).

One fact which seems to support this approach comes from historical linguistics. Many of the words with a ∅ ~ high vowel alternation are loanwords from Arabic and Persian which in their original form had complex codas not otherwise attested in Turkish. For example, the word *adil* [ɑdil] 'justice' undergoes this alternation (the accusative form is *adli* [ɑdli]). This word originally came from Arabic [ɑdl] 'justice', and [dl] coda clusters are not attested in Turkish. So, with this information, it would make sense that the vowel in the second syllable of the bare form is present to repair what would otherwise be an illicit coda cluster.

There are some issues with this approach, however, which I will outline here. The first problem is the data reported by Göksel & Kerslake (2006) in (40) above. It is not clear, if this ∅ ~ high vowel alternation occurs purely to resolve illicit coda clusters, why vowel height of the added suffix should affect this alternation. Regardless of the height of the vowel in the suffix, the coda cluster has been resolved as a result of adding the affix. If the epenthesis account is correct, inserted vowels are not present in the underlying representation, and are added solely to resolve illicit clusters. Thus, a vowel should never surface in the second syllable of the root because the coda cluster has already been resolved and the rule should not apply.

The second and most serious issue, brought up by Bacanlı (2020), is that many coda clusters with inserted vowels *are* actually licit coda clusters in Turkish, and in fact are present in other roots which show no alternation. See below for some of the examples given in Bacanlı (2020:27-28). The relevant coda clusters have been bolded for clarity.

(42) *Comparing coda clusters with inserted vowels to those without (Bacanlı 2020:27-28)*

	Bare Form		*English*	*Suffixed Form*	*English*
a.	kasıt	[kɑ.**sɯt**]	'intention, purpose'	kastım [kas.t-ɯm]	'my purpose'
	ka**st**	[kɑ**st**]	'caste, intent'		
	üst	[y**st**]	'upper plane'		
	do**st**	[do**st**]	'close friend', etc.		
b.	lahit	[lɑ.**hit**]	'sarcophagus'	lahdim [lah.d-im]	'my sarcophagus'
	ta**ht**	[tɑ**ht**]	'throne'		
c.	ufuk	[u.**fuk**]	'horizon'	ufkum [uf.k-um]	'my horizon'
	ze**vk**	[ze**vk**]	'pleasure'		
	zi**ft**	[zi**ft**]	'tar'		
	a**şk**	[ɑ**ʃk**]	'love'		
d.	akis	[ɑ.**kis**]	'echo, reflection'	aksim [ak.s-im]	'my echo'
	ra**ks**	[rɑ**ks**]	'dance'		
	fa**ks**	[fɑ**ks**]	'fax'		
	bo**ks**	[bo**ks**]	'boxing'		
e.	a**kit**	[ɑ**kit**]	'treaty'	aktim [ak.t-im]	'my treaty'
f.	na**kit**	[nɑ**kit**]	'cash'	naktim [nak.t-im]	'my cash'
g.	va**kit**	[vɑ.**kit**]	'time'	vaktim [vak.t-im]	'my time'
h.	sı**kıt**	[sɯ**kɯt**]	'miscarriage'	sıktım [sɯk.t-ɯm]	'my miscarriage'
	se**kt**	[se**kt**]	'sect'		
	a**kt**	[ɑ**kt**]	'act'		
	pa**kt**	[pɑ**kt**]	'pact, treaty'		

In light of these examples (and others not listed here), it becomes impossible to predict, based solely on phonotactics, when a vowel should be inserted into a coda cluster. There is no way to know, for instance, that a vowel must be inserted to break up the /st/ coda cluster in /kɑst/ 'intention, purpose' to derive the correct surface form *kasıt* [kasɯt], but that the same need not be done to resolve the /st/ coda cluster in *üst* /yst/ 'upper plane'. Which roots undergo this alternation *must* to some extent be memorized by speakers- and therefore must be at least partially lexical, rather than purely phonological.

Clements & Sezer (1982) actually acknowledge this problem in a limited way in a footnote, stating that 'there are a few forms in which the epenthetic vowel breaks up clusters which are otherwise permissible in Turkish' (pp 254). Their proposed solution for what they call 'exceptional' words is to argue that they 'are distinguished from normal words in the lexicon in that their final consonants are (unpredictably) extrasyllabic, that is, unaffiliated to the syllable' (pp 254). This idea, that certain words are marked for an inserted vowel by a lexical specification, will inform the account given in Chapter 4. However, rather than arguing that this lexical specification is used to mark only a small number of words which have this alternation, I argue that there is a lexical mechanism that marks *all* words which undergo this alternation.

Because of this issue with unpredictability, Bacanlı et al. (2020) suggest a morphologically specified vowel syncope approach, in line with Erdal (2010). In this approach, roots which undergo this alternation are morphologically specified for vowel syncope when their final consonant is resyllabified into the onset of certain suffixes. With this approach, the data from Bacanlı et al. (2020) is no longer an issue. The word *kasıt* [kasɯt] 'intention, purpose' is underlyingly /kasɯt/, with a vowel in the second syllable that is marked for syncopation. It poses no issue that [st] coda clusters appear in other words, since there is no claim of an illicit cluster or vowel epenthesis. Another word of the same form, such as *kısıt* [kɯsɯt] 'constraint' does not undergo syncope when suffixed (see *kısıtım* [kɯsɯt-ɯm] 'my constraint'), because this word is not marked for syncope and thus keeps the second vowel in the root.

Erdal (2010) argues that inalienability plays an important role in this form of syncope. As mentioned earlier, a high percentage of words for body parts undergo this ∅ ~ high vowel alternation (*omuz* [omuz] 'shoulder', *boyun* [bojun] 'neck', and *burun* [buɾun] 'nose' all undergo this alternation, for example). Body parts are considered inalienable because they are inherently possessed by some person. Most body part words are of Turkic origin, so the argument that they were loanwords which needed a coda cluster repair due to the influence of non-Turkic phonotactics is not available for this set of words. Erdal (2010) argues that these inalienable words developed syncope diachronically because they are so frequently used to with possessive suffixes (possessive suffixes are vowel-initial when following a consonant) and so strongly tied to the person they are possessed by that they form a 'strong juncture' with such suffixes, conditioning the dropping of the second vowel.

Erdal (2010) lays out the idea that certain words are eligible for syncope, while others are not, but does not go into the specifics of how such a distinction may be represented in the lexicon. The idea that certain words are marked for a ∅ ~ high vowel alternation in the second syllable will be further elaborated on in the analysis given in Chapter 4.

2.4.3 Vowel alternation conclusion

This section has given some background on a ∅ ~ high vowel alternation which occurs in the coda clusters of some Turkish words. In addition, two accounts were summarized: the epenthesis account, in which a vowel is inserted to break up an illicit coda cluster, and a syncope account, in which certain words are marked for syncope of the vowel in the final syllable.

2.5 Emphatic Reduplication

This section will cover an array of topics and past research related to Turkish emphatic reduplication (TER). Other forms of Turkish reduplication will be discussed, as well as emphatic reduplication cross-linguistically in Turkic languages. Issues related to the semantics of TER will be briefly covered. Finally, a number of studies focused on the choice of linker in TER will be explained in detail.

2.5.1 Background on Turkish emphatic reduplication

Turkish has several different types of reduplication, one of which is TER. TER is a partial prefixing reduplication applied to derive an emphatic meaning for a limited number of adjectives and adverbs (Göksel & Kerslake, 2006). TER is of particular interest to linguists due to the irregularity of the linker consonant on consonant-initial reduplicated forms. See the below examples of emphatic reduplication in (42). The reduplicated prefix, along with the linker consonant, have been bolded for clarity.

(42) *Examples of Typical Partial Reduplication*

Bare Form	*Translation*	*Reduplicated Form*	*Translation*
güzel [gyzel]	'beautiful'	**güp**güzel [**gyp**gyzel]	'extremely beautiful'
mavi [mɑvi]	'blue'	**mas**mavi [**mɑs**mɑvi]	'very blue, stark blue'
eski [eski]	'old'	**ep**eski [**ep**eski]	'very old'
siyah [sijɑh]	'black'	**sim**siyah [**sim**sijɑh]	'very black, pitch black'
doğru [doːɾu]	'true, towards'	**dos**doğru [**dos**doːɾu]	'honest and aboveboard, dead ahead'
temiz [temiz]	'clean'	**ter**temiz [**ter**temiz]	'neat as a pin'

As shown in (42), reduplicated forms usually consist of the first onset (if present) and mora of a word, plus an additional linker consonant, prefixed on the stem. Note that that the first mora of the base is copied and *not* the first vowel; if the first vowel of the base is long, the reduplicated piece will have a short vowel (see the example of *doğru* [doːɾu] in 42 above)[8]. If the stem begins with a vowel, the linker consonant will be [p] (Göksel & Kerslake 2006). If the stem begins with a consonant, the linker consonant can be one of the following: [p s m ɾ]. Some consonant-initial stems also have more than one possible linker consonant, such as *diri* [diɾi] 'alive, energetic' which can be reduplicated either as *di-**p**-diri* [di-**p**-diɾi] or as *di-**m**-diri* [di-**m**-diɾi] 'full of life'.

In some cases, an optional -A (a low harmonizing vowel which will surface as either [-e] or [-a]) is added alongside the linker consonant (Göksel & Kerslake 2006). See (43) below for examples of these reduplicated forms.

[8] See also Stachowski (2014:213) for a discussion of this pattern for emphatic reduplication in Turkish and other other Turkic languages.

(43) *Partial Reduplication with Extra Vowel*

Bare Form	*Translation*	*Reduplicated Form*	*Translation*
gündüz [gyndyz]	'daytime, by day'	**güp(e)**gündüz [**gyp(e)**gyndyz]	'in broad daylight'
sağlam [sɑːlɑm]	'durable, sturdy'	**sap(a)**sağlam [**sɑp(ɑ)**sɑːlɑm]	'as sound as a bell'

In a few forms where the linker consonant is <r> [ɾ], an additional -Il (I is a harmonizing high vowel which will harmonize in rounding and backness with the preceding vowel) or -Am is added (Göksel & Kerslake 2006).

(44) *Partial Reduplication with Extra Vowel and Consonant*

Bare Form	*Translation*	*Reduplicated Form*	*Translation*
çıplak [tʃɯplɑk]	'naked'	**çırıl**çıplak [**tʃɯɾɯl**tʃɯplɑk]	'stark naked'
parça [pɑɾtʃɑ]	'piece'	**param**parça [**pɑɾɑm**pɑɾtʃɑ]	'torn to shreds, smashed to pieces'

Despite the relatively small number of Turkish words that are able to undergo partial reduplication (and exceptional forms like the ones in 43 and 44), some researchers have argued that native Turkish speakers use phonology to determine the linker consonant when applying TER to novel Turkish words and nonce words (e.g., Taneri 1990, Kelepir 2001, Wedel 1999). The choice of linker consonant is partially predicted by the stem in that the linker consonant is likely to be different from the consonants of the stem. In forms that take TER, the linker consonant is never the same as the first consonant of the stem (Göksel & Kerslake 2006, Kelepir 2001).

2.5.2 Other kinds of reduplication in Turkish

As well as emphatic reduplication, Turkish has two other kinds of reduplication: total reduplication (also referred to as ‘doubling’ in Göksel & Kerslake 2006) and m-reduplication. Total reduplication is a complete duplication of a word, and in Turkish it creates adverbs. For example, the adjective *yavaş* [jɑvɑʃ] ‘slow’ can be fully reduplicated to form the adverb *yavaş yavaş* [jɑvɑʃ jɑvɑʃ] ‘slowly’.

The final form of reduplication, m-reduplication, is a partial suffixing reduplication. In the reduplicated part, the onset of the first syllable (if present) is removed and replaced with [m]. The function of m-reduplication is to add a generalizing meaning, often translated as ‘and the like’ or ‘and stuff’. This reduplication has a generalizing meaning e.g., *çanta* [tʃɑntɑ] ‘purse’ can be reduplicated to *çanta manta* [tʃɑntɑ mɑntɑ] ‘purses and the like’. In addition, m-reduplication cannot be applied to m-initial words, a fact we will return to later.

I will discuss how both of the above forms of Turkish reduplication can be represented with EFP in Chapter 5.

2.5.3 Emphatic reduplication cross-linguistically in Turkic

Partial prefixing emphatic reduplication is present in many Turkic languages, though not all Turkic languages use the same linker consonants. Stachowski (2014) differentiates between what he labels *p-languages*, which seem to exclusively or near-exclusively use [p] as a linker (Bashkir, Dolgan, Khakas, Kirghiz, Karakalpak, Kumyk, Kazakh, Oriot, Shor, Tartar, Tuvinian, Uighur, and Western Karaim) and what he calls *mprs-languages*, which use linkers from the set [p m s r] (Azeri, Eastern Karaim, Gagauz, Ottoman, modern Turkish, Turkmen, & Uzbek). Interestingly, Yakut seems to be an outlier in that it has attested linkers not from [p m s ɾ]. Most attested forms in Yakut have [p] as a linker, but Yakut also uses [s r] as well as linkers not attested in other Turkic languages: [č] (one form), [j] (one form), [k] (one form), [n] (three forms), [ŋ] (one form), and [t] (one form) (Stachowski 2014:184).

The mprs-languages are from the Oghuz group or languages that were heavily influenced by it, which are the westernmost geographically. The p-languages are comprised of the other Turkic languages. Yakut is an outlier which was isolated from other Turkic languages for a long time (Stachowski 2014:216-218).

Turkish appears to be an outlier in that it has the highest number of attested forms and also the most variation in what linker consonant can be chosen for a given base as compared to Turkic languages (Stachowski 2014:134). In my view, it is a bit difficult to gauge if this apparent abundance and variation of emphatic replication in Turkish as compared to other Turkic languages reflects true cross-linguistic variation; rather, it could be a byproduct of Turkish emphatic reduplication having been documented far more in academic literature than that of other Turkic languages. As I will cover in this literature review, TER has been written about quite extensively over the past half century. It is possible that if emphatic reduplication in other mprs-languages were given the same amount of linguistic study as TER, such study would uncover that those languages have a similar amount of attested forms and variation in linker consonant. At this point, it is not possible to know with certainty; in my view this would be an interesting area for further study, though outside of the scope of the present work.

2.5.4 Dialectical variation in TER

The effect of dialectical variation on the choice of linker in real TER forms is understudied. Most studies simply do not specify the dialect of the speakers involved. Taneri (1990), discussed below in section 2.5.7, briefly discusses a possible effect of dialect on TER, but does not offer any further details. Demir (2017), a nonce word study described in more detail in 2.5.11, mentions that dialect was not found to be correlated to choice of linker. Aside from these short mentions, I am aware of only one study which explicitly examines TER in a specific dialect of non-İstanbul Turkish.

Abbasi & Moradkhani (2012) examined emphatic reduplication data from Azerbaijani Turkish. In Azerbaijani Turkish, attested linker consonants are from the set {p m r}. Interestingly, speakers of this dialect never select [s] as the linker consonant in reduplicated forms. Abbasi & Moradkhani (2012) discuss the possibility that, due to its rarity, the linker consonant [ɾ] is fully lexicalized. However, they conclude that the linker consonant is selected from [p] first, and then [m] or [ɾ] based on dissimilation with consonants in the base, though they concede that the rules they posit have many exceptions. Their rules for determining linker are similar to that of other works on Anatolian Turkish which will be discussed further later on. Specifically, these rules are similar to that of Taneri (1990), without [s] as a possible linker consonant (due to it not being present in this dialect) and Wedel (1999), without the lexicalization of the linker consonant [ɾ].

2.5.5 Semantics of TER

The studies in this section investigated TER from a different perspective from the other works on the topic, which primarily focus on the selection of linker. Recall that TER only applies to a limited number of Turkish words, so Turkish speakers must learn which words can undergo this process, and perhaps semantic categories play a role. To my knowledge, two studies have investigated this topic: Kaufman (2014) and Turgay & İskender (2021). Both will be discussed here.

Kaufman (2014) investigated how semantic categories might affect the application of TER by testing Turkish speakers' willingness to apply TER to novel forms of different kinds. The study consisted of two parts: one with nonce forms and one with real Turkish words.

In the first experiment, 45 native Turkish speakers were given several sentences of context which demonstrated that a nonce word fit into one of four semantic classes: color, dimension, texture, and mental state. While color and dimension are semantic categories of words that undergo TER (e.g., ma-**s**-mavi 'extremely blue' and i-**p**-ince 'incredibly thin'), texture and mental state are not. Speakers were then asked to rate on a sliding scale whether they preferred a version of the nonce word intensified with TER (with their choice of linker from the set <p m s r>) or one which was intensified using the adverb *çok* [tʃok] 'very'[9]. Results showed that speakers were more willing to apply TER to nonce color words than words of the other three categories. Interestingly, for dimension, texture, and mental state, results were bimodal, indicating that speakers tended to either strongly prefer or strongly disprefer TER for these categories.

The second experiment had 45 native Turkish speakers complete the same task described above, but with real Turkish words from five different groups. These groups are: Turkish words to which TER normally applies, synonyms of words to which TER normally applies, words from the same semantic class as words to which TER normally applies, gradable adjectives, and ungradable adjectives. The idea was that synonyms and words with the same semantic class, being semantically related to words which receive TER, would be more preferred for reduplication. Finally, there were gradable and ungradable adjectives that were semantically unrelated to words that normally receive TER. The gradable adjectives were expected to be more preferred for reduplication because they are able to be emphasized. However, results only indicated that words to which TER normally applies were more preferred for TER. The other four categories showed no significant difference from each other. Interestingly, boundedness (which was not a purposefully constructed category for either experiment) did have a significant effect, with bounded adjectives being more preferred for TER than unbounded adjectives.

Overall, results of the study showed that semantic class does seem to have some effect on whether speakers are willing to apply TER to a form. For nonce words, color words were significantly more preferred for TER than words from other semantic classes. However, the

[9] Note that TER cannot co-occur with *çok* [tʃok] 'very'. For example, **çok ma-s-mavi* *[tʃok ma-s-mavi] is ungrammatical. So there is no issue here with speakers preferring both.

effect was not apparent in real Turkish words which do not usually undergo TER, indicating that TER is equally dispreferred for words that do not receive TER in the speakers' lexicon.

Turgay & İskender (2021) also investigated the semantics of TER, but from a purely theoretical rather than experimental lens. Turgay & İskender (2021) first point out that TER can only apply to adjectives and adverbs which are gradable. They then make an important distinction that there are two possible meanings that TER can create based on whether the gradable base has a maximal endpoint or a non-maximal endpoint.

Maximal endpoint adjectives have an upper bound. For example, *clean* is a maximal endpoint adjective because there is a limit to how clean something can be: once an object is completely clean, it cannot get any cleaner. Turgay & İskender (2021) point out that when TER is applied to maximal endpoint adjectives and adverbs, the emphasized meaning created by TER is 'completely'. For example, *temiz* [temiz] 'clean' can undergo TER to form *tertemiz* [tertemiz], which can be translated as 'completely clean'. Non-maximal endpoint adjectives and adverbs do not have any upper bound. For example, long is a non-maximal endpoint adjective, because even if an item is very long, it can be even longer. Turgay & İskender (2021) argue that when TER is applied to a non-maximal endpoint base, it yields a 'very' meaning. For example, *uzun* [uzun] 'long' can undergo TER to form *upuzun* [upuzun] which yields the meaning 'very long'.

I disagree with Turgay & İskender (2021)'s conclusion that which the set of words which can undergo TER is predictable by semantic class, as expressed in the quote below (from pp 120):

> 'We thus conclude, contra the literature, that partial reduplication in Turkish is not an unpredictable lexical process. The group of adjectives that participate in partial reduplication, and the resulting interpretation when they do so, is predictable by their semantic class.'

It is true that only gradable bases can undergo TER. However, not every gradable adjective and adverb is allowed to undergo TER. For example, Kaufman (2014) identified *dev* [dev] 'giant' as a synonym of *koca* [kodʒɑ] 'huge'. The word *koca* can be reduplicated to *koskoca* [koskodʒɑ] 'extremely huge'. However, *dev* [dev] 'giant' is not a word which can undergo TER (Kaufman 2014). Kaufman (2014) found in her second experiment (see above) that speakers were *not* more willing to reduplicate such synonyms than other adjectives, even compared to *nongradable* adjectives. So, which words can undergo TER is partially, but not completely, predictable based on their semantic class. Thus, speakers must somehow memorize or mark words in their lexicon which are allowed to undergo TER, to differentiate this partially arbitrary set from words which cannot undergo TER.

The semantic studies discussed above (Kaufman 2014 and İskender & Turgay 2021) are less related to the current work than the other studies I will discuss in the preceding subsections, in that they do not address the selection of the linker consonant, nor are they directly related to phonology at all. However, they investigate a key detail regarding TER which will be discussed in this dissertation: since TER only applies to certain words, how do speakers determine which words can undergo TER? I have argued that this must be marked somehow in the lexicon. Exactly how this marking can be represented using EFP will be covered in Chapter 5.

2.5.6 Lexicalized explanation for TER linker selection

Some grammars of Turkish have characterized the choice of linker as partially arbitrary (Underhill 1976, Göksel & Kerslake 2006) or completely arbitrary (Swift 1963, Lewis 1967, Underhill 1976, van Schaaik 2020). In these accounts, the choice of linker is not determined by phonological rules- linker consonants are lexicalized and memorized with the base. Other works which argue against the choice linker being phonological include Yılmaz (2020), Kılıç & Bozşahin (2013) and Demir (2013, 2018). This stance can be summed up well by the following quote from Swift's (1963) grammar of Turkish, also highlighted in Taneri (1990):

> 'If the word commences with a consonant, /p/, /m/, /r/ or /s/ may occur. *There does not seem to be any discernible phonological conditioning of the choice of one rather than another of these and alternate forms do exist for some words with different consonants*' (Swift 1963:123, italics added)

However, a number of phonologists have posited that the choice of linker consonant can be determined by phonological rules, either through identity avoidance, or via dissimilation with consonants of the base (e.g.,Taneri 1990, Kelepir 1999). This has spawned a great volume of research on TER aimed at understanding the mechanisms for linker selection, summarized in the following subsections.

2.5.7 Studies of real Turkish forms that undergo TER

Demircan (1987, as cited in Demir 2017) presented experimental data pertaining to the phonology of emphatic reduplication in Turkish. In the study, native Turkish speakers chose whether the emphatic form of a word would have [p], [m], [ɾ], or [s] as its linker consonant. The study included actual Turkish adjectives and adverbs, as well as nonce words. Results indicated that the underlying form for the linker consonant is /p/, which is changed to one of [m] > [ɾ] > [s] (the priority goes in that order) based on dissimilation with the consonants in the base.

Taneri (1990), another experimental study, elicited data from 32 native Turkish speakers about the reduplicated forms of approximately 300 Turkish adjectives and adverbs. Of those approximately 300, ~280 had an attested reduplicated form from at least one speaker. The remaining twenty-odd forms could not be reduplicated by any of the 32 participants. The design of the experiment has a methodological issue which makes interpreting the results a bit difficult. Recall that TER only applies to a limited set of words. Since speakers were given a very large list of words and asked to apply TER to each of them, it is unclear whether, for each word, a speaker was applying TER to a word they would usually apply it to, or if they were treating it as a novel word and attempting to expand their class of words which undergo TER. Based on further research conducted in the intervening interval since Taneri's (1990) study, it seems that these tasks yield different results in terms of linker selection, which will be explained further in coming subsections.

Turning to the study results, Taneri (1990) observed large variations in how generously speakers applied TER to the words in the study. While some speakers were able to reduplicate many words, other speakers were much more conservative in their application of TER. The initial intent of the study was to examine differences in TER based on sex and age, but no differences were found with regard to these variables. Rather, Taneri states that the differences between subjects with regard to how they applied TER were likely largely due to register. Participants who felt comfortable with the study and were interested in language were much more likely to apply TER many words. Others who felt less comfortable, however, were much less generous in their application of TER. Taneri also speculates that some inter-speaker variation may be due to regional differences, but does not elaborate, because the participant pool was not constructed to examine those differences.

In addition to the inter-speaker differences in number of words reduplicated, Taneri found that many adjectives and adverbs had more than one attested linker consonant. For example, Taneri (1990) reports two possible reduplications of doğru [doːɾu] 'towards': *do**p**doğru* [do-**p**-doːɾu] (26 speakers), as well as *do**s**doğru* [do-**s**-doːɾu] (6 speakers), meaning 'honest and aboveboard, dead ahead'. Taneri reported that in some cases speakers were aware of multiple accepted linkers for a given word, and that they would accept any of those known linkers as correct. Sometimes the addition of a different linker consonant created a different meaning. For instance, some speakers indicated that *ka**s**katı* [ka-**s**-katɯ] was used to mean 'very stiff', while *ka**p**katı* [kɑ-**p**-kɑtɯ] meant 'very thick' (as a liquid).

The findings of the study included a hierarchy in the choice of linker consonant, which was [p] > [s] > [m] > [ɾ] based on how frequently each was used, consistent with Demircan (1987). Taneri (1990) posited that the choice of linker consonant is the result of dissimilation (in terms of both place and manner of articulation) with consonants in the base, primarily with the first two consonants. However, Taneri (1990) notes that there are forms which take variable linkers that her rules cannot account for. Each of the rules posited in Taneri (1990) is outlined below and data is given which conflicts with them. The purpose here is not to disprove the analysis laid out in Taneri (1990) (as the author herself states that the rules cannot encompass all of the data); rather, the goal is to highlight the difficulties encountered when creating rules to account for the selection of the linker in TER.

First, Taneri notes that the linker [p] was never the sole selected linker for words containing the bilabials [p b m], so she posits a rule which blocks [p] from being selected as a linker for words with a bilabial consonant. Taneri is correct in that there seem to be very few words which *solely* take [p] as a linker which contain a bilabial consonant. There are a number of words, however, which contain [b m] that can take [p] as a one possible linker. I have provided a table of twelve such forms below, along with the source or sources in which a [p] linker is reported, and alternate linker or linkers which have been reported for the word.

(45) *TER forms with [p] linker that also have a bilabial consonant*

TER form	*IPA*	*Translation*	*Source(s) which attest [p] linker*	*Alternate linker(s)*
cıpcı**b**ıl	[dʒɯ-p-dʒɯ**b**ɯl]	stark naked	Özçelik (2012), Yılmaz (2020)	s
cıpcı**b**ıldak	[dʒɯ-p-dʒɯ**b**ɯldɑk]	stark naked	Özçelik (2012), Yılmaz (2020)	s
çapça**b**uk	[tʃɑ-p-tʃɑ**b**uk]	very fast	Wedel (1999), Özçelik (2012), Yılmaz (2020)	m, r, s
kıpkır**m**ızı	[kɯ-p-kɯɾ**m**ɯzɯ]	extremely red	Wedel (1999), G & K (2006), Özçelik (2012), Stachowski (2014), Köylü (2021), Schaaik (2020)	s
köpkötürü**m**	[kœ-p-kœtyɾy**m**]	completely immobile	Özçelik (2012)	s
tapta**m**a**m**	[tɑ-p-tɑ**m**ɑ**m**]	complete	Özçelik (2012), Yılmaz (2020)	m (1 speaker from Taneri 1990), s
tepte**m**iz	[te-p-te**m**iz]	completely clean	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	r
sapsağla**m**	[sɑ-p-sɑːlɑ**m**]	healthy	Swift (1963), Wedel (1999), G & K (2006), Ketrez (2012), Özçelik (2012), Yılmaz (2020)	alternate form with additional vowel: sapasağlam
yapya**b**ancı	[jɑ-p-jɑ**b**ɑndʒɯ]	very foreign	Stachowski (2014), Yılmaz (2020)	-
yapya**m**alak	[jɑ-p-jɑ**m**ɑlɑk]	very sloppy	Yılmaz (2020)	s
yupyu**m**ru	[ju-p-ju**m**ɾu]	completely round	Yılmaz (2020)	s
yupyu**m**uşak	[ju-p-ju**m**uʃɑk]	very soft	Taneri (1990), Özçelik (2012), Yılmaz (2020), Köylü (2021)	m, r, s

Taneri (1990) posits a further rule blocking the selection of [s] for words with a sibilant in C1 position. Taneri (1990) does not clarify exactly what set of consonants comprise sibilants in Turkish, so it is unclear if the affricates *ç c* [tʃ dʒ] are meant to be included in this generalization. If the affricates are included in the sibilant class, this creates a problem, as there are a number of affricate-initial words which can take *s* as a linker. Several examples are shown below:

(46) *Affricate-initial words which can take [s] linker*

TER form	*IPA*	*Translation*	*Source(s) which attest [s] linker*
cascavlak	[dʒɑ-s-dʒɑvlɑk]	completely bald	Wedel (1999), Kelepir (2001), Göksel & Kerslake (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)
cıscıvık	[dʒɯ-s-dʒɯvɯk]	extremely sloppy	Wedel (1999), Kelepir (2001), Göksel & Kerslake (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)
çasçabuk	[tʃɑ-s-tʃɑbuk]	very fast	Özçelik (2012), Yılmaz (2020)
çısçıplak	[tʃɯ-s-tʃɯplɑk]	stark naked	Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)
çesçevre	[tʃe-s-tʃevre]	all around	Yılmaz 2020

If affricates are excluded from the sibilants, then I agree with Taneri (1990) that for real TER forms, there are no attested words with a sibilant as C1 which take [s] as a linker. However, there are some important details to note here. First, there are four sibilant fricatives in Turkish: *s z ş j* [s z ʃ ʒ]. In Turkish, C1 cannot match the linker, so [s] is never chosen for [s]-initial forms. There are no attested TER forms beginning with [ʒ], since this is a loan from

French with low prevalence in Turkish in general. This leaves *ş z* [ʃ z], of which there are five attested forms that I can find (with several more found only in Taneri 1990)[10]. For these forms, [s] is never attested as being chosen as the linker consonant. The relevant forms are given in (47) below.

(47) *Attested s- and z- initial TER forms*

TER form	*IPA*	*Translation*	*Source(s)*
şepşekerli	[ʃe-p-ʃekeɾli]	extremely sugary	Wedel (1999), Taneri (1990)
şemşekerli	[ʃe-m-ʃekeɾli]	extremely sugary	Taneri (1990) (1 speaker)
şipşirin	[ʃi-p-ʃiɾin]	extremely cute	Taneri (1990), Wedel (1999), Kelepir (2001), Göksel & Kerslake (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)
zapzayıf	[zɑ-p-zɑjɯf]	thin as a rail	Swift (1963), Wedel (1999), Kelepir (2001), Göksel & Kerslake (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)
zepzengin	[ze-p-zengin]	very wealthy	Wedel (1999), Yılmaz (2020)
zopzor	[zo-p-zoɾ]	incredibly difficult	Stachowski (2014), Yılmaz (2020)

Given how few words beginning with sibilant fricatives can undergo TER at all, it perhaps seems less surprising that there are none which take [s] as the linker. In addition, [p]

[10] Wedel (1999) attests the form *şipşirkin* [ʃi-p-ʃiɾkin]. I cannot find the words *şirkin* or *şipşirkin* in Turkish dictionaries. I believe in may be a mistranscription of *çipçirkin* [tʃi-p-tʃiɾkin] 'extremely ugly', so I excluded it from the table below.

is the most common linker (Taneri 1990, Stachowski 2014). In novel and nonce word studies (described in detail in sections 2.5.8 and 2.5.11), [s] seems to be dispreferred for linkers starting with *ş z* [ʃ z], but some speakers do still select it (Demir 2017, Kılıç & Bozşahin 2013, Köylü 2021).

Thirdly, Taneri (1990) posited a rule blocking the selection of [m] for words that contain nasals in any position, a bilabial as the second consonant, or [l ɾ] as either of the first two consonants. See below for a table of words which take [m] as a linker but do not conform to this rule:

(48) *Words which can take [m] linker not captured by Taneri (1990)*

TER form	*IPA*	*Translation*	*Source(s) which attest [m] linker*
bumbu**lan**ık	[bu-m-bu**lan**ɯk]	completely blurry	Stachowski (2014), Yılmaz (2020)
yumyu**m**uşak	[yu-m-yu**m**uʃɑk]	extremely soft	Taneri (1990), Yılmaz (2020)
bombo**l**	[bo-m-bo**l**]	extremely abundant	Taneri (1990), Yılmaz (2020)
cımcı**lk**	[dʒɯ-m-dʒɯ**lk**]	completely rotten	Wedel (1999), Özçelik (2012), Yılmaz (2020), Stachowski (2014)
bembe**r**k	[bembe**r**k]	extremely strong	Yılmaz (2020)
pempe**r**iş**an**	[pe-m-pe**r**i**ɑn**]	completely miserable	Demir (2017)

Finally, Taneri posits a rule which blocks the selection of *r* [ɾ] as a linker when *r* [ɾ] is present in the base. There are two attested forms with an *r* [ɾ] present in the base, which have been reported to take an *r* [ɾ] linker, shown in the table below.

(49) *Bases which can take [ɾ] that contain [ɾ]*

TER form	*IPA*	*Translation*	*Source(s) which attest [ɾ] linker*
hırhışı**r**	[hɯ-ɾ-hɯʃɯ**ɾ**]	extremely stupid	Yılmaz (2020)
tortopa**r**lak	[to-ɾ-topɑ**ɾ**lɑk]	very round	Özçelik (2012)

In conclusion, while the rules laid out by Taneri (1990) can capture general trends in the choice of linker, there is some data that cannot be captured in this manner. As mentioned earlier, Taneri's (1990) paper even recognizes this issue, calling for further research to understand the exact mechanism behind linker selection. This typifies the general difficulty with crafting rules to fully predict the choice of linker consonant: while it is possible to capture loose patterns, there are always exceptions which do not take the predicted linker.

There is one more study on attested TER forms which takes a quite different approach from Demircan (1987) and Taneri (1990). Tang & Akkuş (2023) ran an acceptability rating task of 162 real consonant-initial base forms with 207 total native Turkish speakers. The study and its results will be summarized here first, followed by an explanation of an important methodological issue below. In the experiment, speakers were shown words TER applied, and a different linker consonant from the set [p m s ɾ]. Each speaker saw about 30 forms, and every form was rated by at least 40 speakers. Speakers were asked to rate each linker consonant from 1 DOĞAL DEĞİL 'not natural' to 7 DOĞAL 'natural'. The goal of the study was to investigate whether the selection of the linker was influenced by partial identity avoidance with individual features (e.g., a dispreference for a strident adjacent to another strident) or total identity avoidance (e.g., a dispreference for an [s] adjacent to another [s]). In addition, the study looked at locality effects, i.e., whether consonants closer to the beginning of the word influence the choice of linker more than consonants later in the word.

Results were analyzed using mixed effects linear regression, and indicated that both total identity avoidance and dissimilation with specific features influenced acceptability ratings, though total identity was a more important factor. In addition, the statistical model indicated a gradient effect, where consonants closer to the beginning of the base were more important for the rating of a linker. However, syllable structure also seemed to play an important role, with codas contributing more than onsets. This sometimes reversed the overall trend, for example, in CVCCV words, the coda C2 was more important than the onset C1.

It is very important to point out here that there is a methodological issue with the statistical analysis performed in Tang & Akkuş (2018). The type of data collected, which consists of subjective judgments on a scale from 1 to 7 (in increments of whole numbers) is not well-suited to a mixed effects linear regression model. Linear regression is appropriate only for data which is continuous and has consistent distances between values, and the data collected is neither. Firstly, the rating data is not continuous, and ratings from 1 to 7 are in seven discrete categories[11]. Secondly, the distance between rating increments is not necessarily consistent, because these are subjective judgments. For example, the speakers' understanding of the magnitude of difference between a rating of 0 and a rating of 1 may not be the same as their understanding of the magnitude of difference between a rating of 4 and a rating of 5. Due to the statistical analysis being incompatible with the data collected in the study, it is difficult to draw any conclusions from the results of the study.

The rating data collected Tang & Akkuş (2018) would be much better suited to an ordinal mixed effects regression model. Ordinal regression is well-suited for rating data of this type, which is in discrete categories that are ordered, but may not have consistent spacing between the categories. The data from this study is publicly available for download, and I encourage the reader to retrieve and re-analyze the data in this way.

[11] While it is not entirely clear in the text of the paper itself that the ratings collected were in whole number increments rather than a continuous sliding scale, I looked at the raw data from the study which is posted publicly online and confirmed that the ratings were indeed only whole numbers.

2.5.8 Experimental studies with real novel forms

Wedel (1999) combines data about real attested forms with experimental data on speakers applying TER to novel forms and arrives at a mixed lexical and dissimilation approach to determine the linker consonant. The experiment described in the study consisted of three to nine native Turkish speakers (number of responses varied by word) applying emphatic reduplication to approximately 65 Turkish adjectives to which it could not normally be applied. These new adjectives were included as a way to solicit novel reduplicated forms. All speakers used a linker consonant from the set [p m s], though only about half of the 65 words were reduplicated with the same linker consonant across all speakers. For example, *bezgin* [bezgin] 'discouraged' was reduplicated by nine speakers. Five speakers produced *be-**m**-bezgin*, three speakers produced *be-**s**-bezgin*, and one speaker produced *be-**p**-bezgin*.

In addition to the experimental data, Wedel (1999) compiled a list of about 125 real attested forms to analyze. Wedel noticed that, in general, speakers dispreferred linker consonants that were identical to C1 or C2 of the base. In addition, speakers never selected [ɾ] as the linker for novel forms. From this data, Wedel (1999) argues that while the linker consonant [ɾ] is completely lexicalized, the choice of linker consonant on a novel word can be [p m s], where [p] is the default linker consonant and a choice of [m] or [s] is determined by dissimilation from the first and second consonant in the base.

There are some issues, however, with Wedel's proposed constraints for dissimilation. In order to represent dissimilation from the first and second consonant, Wedel posits a high ranking *Repeat markedness constraint which assigns a violation if the linker consonant is identical to C1 or C2. Because this constraint outranks all others relevant for selecting the linker consonant, this constraint should prevent the linker from matching C1 or C2 altogether. While it was never the majority of speakers, some speakers in Wedel's data did choose C1 or

C2 as the linker consonant for certain words. See below for a table of these tokens from Wedel's dataset:

(50) *C1 or C2 as linker in novel forms (from Wedel 1999)*

Word	*IPA*	*Translation*	*Reduplicated Form*	*# Speakers*
miskin	[miskin]	sluggish	mi-**s**-mi**s**kin	1 out of 5
pasaklı	[pɑsɑklɯ]	dowdy	pa-**s**-pa**s**aklı	1 out of 5
somut	[somut]	tangible	so-**m**-so**m**ut	1 out of 7
topal	[topɑl]	lame	to-**p**-to**p**al	2 out of 8
müzevir	[myzevir]	troublemaking	mü-**m**-**m**üzevir	1 out of 5

While Wedel's constraints capture trends in the solicited data, they cannot account for the minority of speakers who provided these reduplicated forms with C1 or C2 as the linker consonant. In addition to the novel form results, there are several attested forms which take a linker consonant which is identical to C2 of the base.

(51) *C2 as linker in attested forms*

Word	*IPA*	*Translation*	*Reduplicated Form*	*Source(s)*
yapışık	[japɯʃɯk]	attached	ya-**p**-ya**p**ışık	Özçelik (2012) Yılmaz (2020)
kapalı	[kapalɯ]	closed	ka-**p**-ka**p**alı	Wedel (1999)
yumuşak	[jumuʃɑk]	soft	yu-**m**-yumuşak	Yılmaz (2020) Taneri (1990)
perişan	[periʃɑn]	miserable	pe-**r**-pe**r**işan	Wedel (1999) Kelepir (2001) Göksel & Kerslake (2006) Stachowski (2014) Demir (2017) Yılmaz (2020)

The data above clearly indicate that the constraints posed by Wedel (1999), while capturing trends in the data, cannot account for all of the data, both in novel forms and in real TER forms.

Kılıç & Bozşahin (2013) also conducted a study in which Turkish speakers were asked to apply TER to novel Turkish words. The study had 50 native Turkish speakers reduplicate 31 nouns and verbs. Nouns and verbs were chosen because they are sure to be novel; in natural language, TER only applies to adjectives and adverbs. Of the 31 words included, 30 were consonant-initial, and one word was vowel-initial. As in Wedel (1999), speakers reduplicated the same words in different ways. In fact, all 30 consonant-initial words included in the study were reduplicated with more than one linker consonant. The consonant-initial word which was the most consistent across speakers was *hüzün* [hyzyn] 'sadness' which was reduplicated with [p] by 43 speakers, [s] by 5 speakers, and [m] by 2 speakers.

In addition, results showed that participants reduplicated novel forms in a way that is not attested in real forms. Firstly, in some cases speakers chose a linker which was identical to C1: *masal* [mɑsɑl] 'fable' which was reduplicated as *ma-**m**-masal* (by 5 speakers) and *pırasa* [pɯɾɑsɑ] 'leek' which was redplicated as *pı-**p**-pırasa* (by 14 speakers) (Kılıç & Bozşahin 2013:2725). In attested forms, the linker consonant is never identical to C1. Secondly, in some cases, speakers chose [p] as a linker for b-initial words. In attested forms, [p] is never chosen as a linker when the stem begins with [b]. The two b-initial words used in the study were both reduplicated with [p]: *bıçak* [bɯtʃɑk] 'knife' (reduplicated as *bı-**p**-bıçak* by 11 speakers) and *böcek* [bœdʒek] 'bug' (reduplicated as *bö-**p**-böcek* by 13 speakers).

As in the experimental results for Wedel (1999), C2 was used as a linker for novel forms. In fact, all seven words included in the study which contained a possible linker in C2 position were reduplicated with C2 as the linker by at least one speaker. This part of the results is summarized in the table below, in order from greatest to least number of speakers. No words included in the experiment had [p] in C2 position.

(52) *C2 as linker in novel forms (from Kılıç & Bozşahin 2013:2725)*

Word	*IPA*	*Translation*	*Reduplicated Form*	*# Speakers*
masal	[mɑsɑl]	fable	ma-**s**-ma**s**al	29
pırasa	[pɯɾɑsɑ]	leek	pı-**r**-pı**r**asa	10
cami	[dʃɑmi]	mosque	ca-**m**-ca**m**i	8
resim	[ɾesim]	picture	re-**s**-re**s**im	8
kemir	[kemiɾ]	gnaw	ke-**m**-ke**m**ir	5
şerit	[ʃeɾit]	tape	şe-**r**-şe**r**it	5
surat	[suɾɑt]	face	su-**r**-su**r**at	1

Finally, Kılıç & Bozşahin (2013) cross-referenced their results with consonant-consonant sequences in Turkish in general, using two large corpora (METU-Sabancı Turkish Treebank and BOUN corpus). For each word, they compared the frequency ranking of the consonant-consonant sequence created by the selected linkers and first consonant of the base to the frequency ranking of those same consonant-consonant sequences in the corpora. For instance, they compared the preferences for linkers on b-initial words to the relative frequencies of [pb], [mb], [sb], and [rb] in real Turkish words contained in the corpora. In every case, they found that the frequency ranking for linkers in the study were exactly the *opposite* to the frequencies of consonant-consonant sequences in real Turkish words.

From these results, Kılıç & Bozşahin (2013) argue that the selection of the linker in TER is not phonological, but rather morpholexical. In their view, speakers are relying on global lexical knowledge (frequencies of consonant-consonant sequences in Turkish words at large) to select linkers which result in infrequently attested consonant-consonant sequences.

In addition, the novel words did not conform to the same patterns as attested forms in that C1 could be selected as a linker and [p] was selected as a linker for [b] initial words.

2.5.9 Theoretical phonological treatments

Kelepir (2001) provided a phonological account of linker selection using Optimality Theory (OT). Kelepir (2001) argued for a number of OT markedness constraints which influence the distribution of the linker consonant. These constraints assign violations if the linker consonant matches the first or second consonant of the base in place of articulation, continuancy, or sonority. Below is an example tableau from Kelepir (2001) for the base *katı* [kɑtɯ] 'hard, stiff'. The constraint *αCONT~αCONT assigns a violation when the linker and second consonant of the base match for continuancy, and the *COR~COR constraint assigns a violation when the linker and second consonant of the base are both coronals.

(53) *Tableau for the form kaskatı [ka-s-katɯ] 'extremely stiff'*

(10) /RED+katı/ 'hard'	*αCONT~αCONT	*COR~COR
a) kap.katɨ	*! p-t	✓
b) kam.katɨ	*! m-t	✓
c) ☛kas.katɨ	✓	*

Since the candidates *kapkatı* [ka-p-katɯ] and *kamkatı* [ka-m-katɯ] both violate the highest ranking *αCONT~αCONT constraint, *kaskatı* [kɑ-s-kɑtɯ] is the optimal candidate that appears as the surface form.

Since the tableaux as given select for only one optimal linker candidate per form, Kelepir (2001) cannot account for TER forms which can take more than one possible consonant. For example, the tableau above unambiguously selected [s] as the linker for *katı* [kɑtɯ] 'hard, stiff'. However, [p] is also attested as a linker for this form (Taneri 1990, Stachowski 2014, Yılmaz 2020, Köylü 2021). The constraints as posited by Kelepir (2001) cannot account for this variation.

Additionally, there are some cases in which Kelepir's constraints predict the selection of an unattested linker for a given work. For instance, as Köylü (2021) points out that using the constraints from Kelepir (2001) derive the incorrect reduplicated form of [kodʒa] 'large'. Due to a high ranking markedness constraint against repeated stridents, the incorrect form **korkoca* *[ko-**r-**kodʒa] is selected over the correct *koskoca* [ko-**s-**kodʒa] 'enormous'. Finally, Kelepir (2001) does not put forth any argument concerning the highly exceptional forms in which more than one linker consonant (or an additional vowel) are added.

Özçelik (2012) takes a different approach, forming what is essentially a decision tree for selecting the linker consonant based on the consonants of the base. For example, Özçelik states that if a word begins with a labial obstruent, [p] is not available as a linker. The first choice of linker in this case is [s]. However, if the word contains <s ş z> [s ʃ z] then [s] is excluded as the linker and [m] must be selected. Yılmaz (2020) points out some issues with this approach. For instance, he lists some forms beginning with [b] that are reduplicated with [s], despite not containing <s ş z> [s ʃ z], such as *bombok* [bo-m-bok] 'crappy'. In addition, Yılmaz lists forms containing <s ş z> [s ʃ z] which can take [s] as a linker, such as *yusyumuşak* [ju-s-jumuʃak] 'extremely soft' (Yılmaz 2020:121).

2.5.10 TER in children

Sofu (2005) observed how native Turkish speaking children (aged two through six) used several different kinds of reduplication, including TER. For the TER part of the study, 25 six year old children were asked to reduplicate 38 Turkish nonce words. The same task was given to 25 adults in order to compare across the two groups. In general, the children selected [p], the most common linker consonant on attested forms, more frequently than adults. In addition, children tended to pick consonants which are not typically allowed as linkers, such as [t] or [f], at a higher rate than adults. Adults also sometimes picked linkers not from the set [p m s ɾ], but less frequently. More nonce word studies will be discussed in the preceding sections.

2.5.11 Open ended nonce word experiments

Two more recent studies of TER, Demir (2017)[12] and Köylü (2021), have been focused on open-ended activities with nonce words. Both will be summarized briefly here.

Demir (2017) conducted a study in which 125 native Turkish speakers were given a list of 10 actual Turkish words and 34 nonce words and allowed to reduplicate them in a free response task. Results of the experiment revealed that participants reduplicated nonce words with a wide variety of linker consonants, including copying a consonant from the base as a linker consonant, adding a seemingly random linker consonant, or simply omitting the linker consonant altogether.

Demir (2017) conducted a further followup experiment with 38 native Turkish speakers. The aim of this additional study was to investigate whether nonce words that contained the identical consonants as actual Turkish adjectives and adverbs would be reduplicated with the same linker consonant. This was not the case. For six of the ten nonce words, the predicted linker consonant (the linker from the real word with identical consonants) was selected less than 50 percent of the time. For example, the nonce form *boyuz* [bojuz] has the same consonants as the real Turkish word *beyaz* [bejaz] 'white' which can be reduplicated to *be-m-beyaz* [bembejaz] 'extremely white'. In the study, only 14 speakers chose [m] as the linker consonant for *boyuz* [bojuz].

Due to both the wide range of linker consonants applied to nonce words, and the fact that same-consonant nonce words do not consistently receive the same linker as their real counterparts, Demir (2017) concluded that the choice of linker consonant on actual Turkish

[12] The results on Demir (2017) are also described much more briefly in a proceedings paper, Demir (2018). See Demir (2018) for a concise summary of the experiment and its results, and see Demir (2017) for much more detail on the experiment.

words is likely lexicalized, but that their distribution was originally influenced by phonological factors. Demir (2017) also pointed out that speakers had a strong tendency to avoid picking a linker consonant that was identical to the first consonant of the base.

Köylü (2021) is another recent study on emphatic reduplication which uses a similar approach to Demir (2017). Köylü (2021) asked 14 native Turkish speakers to reduplicate 48 nonce words with various syllable structures (VCV, CVC, CVCV, and VCCV). Consistent with prior findings, Köylü (2021) found that there were a wide variety of linker consonants used to reduplicate nonce words. For example, the nonce form <zımı> was reduplicated as zı-m-zımı (8 speakers), zı-p-zımı (4 speakers), and zı-s-zımı (2 speakers)[13]. However, unlike in Demir (2017), speakers always selected one of [p m s] as the linker consonant. Additionally, Köylü (2021) speculated that in addition to consonants, the vowels of the base may influence the choice of linker consonant.

While choice of linker consonant was variable, Köylü (2021) found that speakers tended to avoid linkers identical to C2, and never picked linkers identical to C1. As a result, Köylü (2021) argued that the dissimilation process observed in Turkish emphatic reduplication is part of a broader cross linguistic tendency away from adjacent segments that are similar to one another, specifically that this is an example of identity avoidance, like Yip (1998)'s identity avoidance principle. Yip (1998) proposed this identity avoidance principle in Optimality Theory to capture the fact that sequences of homophonous phonemes or morphemes are dispreferred cross linguistically. For example, English does not allow both the plural and possessive -s to be pronounced on the same root (demonstrated in 54 below).

[13] Note that Köylü (2021) only reports percentages. I arrived at these speaker totals by comparing the percentages to the total number of speakers (fourteen).

(54) *Identity avoidance principle in English*

cats = [k^{h}æt^{h}s]   cat's = [k^{h}æt^{h}s]

cf. cats' = [k^{h}æt^{h}s]   *[k^{h}æt^{h}s:] or *[k^{h}æt^{h}səz]

In 54 above, the phonologically identical English suffixes [-s] (plural) and [-s] (genitive) are blocked from being simultaneously pronounced the same root [kæt]. Only one [-s] suffix may be phonologically overt. That is, one [-s] is deleted in the surface form. Köylü (2021) argues that the reason for the dissimilation of Turkish linker consonant from the base is this cross-linguistic bias against adjacent homophonous phonemes or morphemes. This idea of identity avoidance is discussed more in Chapter 5, wherein I discuss my own analysis of TER.

2.5.12 A systematic comparison of attested TER forms

Many of the sources cited above contain lists of attested TER forms. However, the data provided overlapped between sources to different degrees, and sometimes sources attributed conflicting linker consonants to the same form. In order to systematically investigate variation in the choice of linker for real forms, I compiled a list of all the attested consonant-initial forms across eleven sources: Swift (1963), Wedel (1999), Kelepir (2001), Göksel & Kerslake (2006), Özçelik (2012), Ketrez (2012), Stachowski (2014), Demir (2017), van Schaaik (2020), Yılmaz (2020), and Köylü (2021). Data from Taneri (1990) was also used as a comparison for words found in at least one of the other sources. Crucially, this list does not contain nonce forms or novel real word experiments; the purpose of the list is to examine data on real, attested TER forms[14]. The full spreadsheet is in Appendix 2 and also available online at this link:

[14] This requirement for only real, attested forms is why data from Taneri (1990) is included only where a word from that study was attested in another source. Based on the construction of the study, it is not possible to tell where speakers were applying TER to a word which they knew was eligible to undergo TER, and where they might have been adding TER to a novel word.

https://osf.io/f4yec/overview?view_only=8a08adc3ab7d405ba7e784a718b6705e

The purpose of the lists of attested forms in the surveyed sources varied. Some lists were brief and served to either give examples to contextualize research (Köylü 2021) or to provide a short guide on the TER process for Turkish learners (Swift 1963, Ketrez 2012, Schaaik 2020). Others served as the data from which linguistic generalizations were made (Wedel 1999, Kelepir 2001, Göksel & Kerslake 2006 , Özçelik 2012, Yılmaz 2020). Finally, some report results from a number of speakers reduplicating real words (Taneri 1990, Demir 2017). Some sources reported multiple linkers where the authors felt it was appropriate, while others only ever chose a single form to report. See below for a table of all the sources, their number of TER words reported, and whether they allowed for multiple possible linkers:

(55) *Sources included in the compiled list of attested TER forms*

Source	*# consonant initial TER forms*	*Includes multiple possible linkers?*[15]
Swift (1963)	20	No
Taneri (1990)	73	Yes
Wedel (1999)	125	Yes
Kelepir (2001)	86	No
Göksel & Kerslake (2006)	78	No
Özçelik (2012)	134	Yes
Ketrez (2012)	39	No
Stachowski (2014)	160	Yes
Demir (2017)	18	Yes

[15] I am excluding here cases where an additional vowel or additional vowel and consonant may be added, such as *güp(e)gündüz* [gyp(e)gyndyz] ‘in broad daylight’. I include only cases where the author allows multiple options from the linker set <p m s r>.

Source	*# consonant initial TER forms*	*Includes multiple possible linkers?*[15]
van Schaaik (2020)	20	No
Yılmaz (2020)	190	Yes
Köylü (2021)	30	Yes

The total number of unique words included in this compiled list was 202. Some words are attested much more than others. About 20% of words (43 total) were attested as undergoing TER in only one source. A further ~15% of words (29 total) were attested as undergoing TER in two sources. This leaves a final ~65% (130 words) which are attested as undergoing TER in at least three different sources.

For words which had only one possible linker attested, [p] was the most common by a wide margin, followed by [s] and [m], which were similarly common, and finally <r> [ɾ] was very uncommon.

(56) *Summary of linkers for words with only one attested linker*

Linker	*# of TER forms*
p	64
m	17
s	26
r	6

There is a fairly high degree of variability in the choice of linker for the words. Approximately 45% of words (89 total) had at least two attested linker consonants. Thirteen words had three different attested linkers. Two words were attested as having four possible linker consonants. The first, *yumuşak* [jumuʃɑk] 'soft', was attested as taking all four linkers in one source (Yılmaz 2020), though other sources reported that it took [p m s] (Taneri 1990), just [p s] (Özçelik 2012, Köylü 2021) or only [s] (Wedel 1999, Stachowski 2014, see also Kelepir 2001, though Kelepir lists only one linker for each word). The second, *çabuk* [tʃɑbuk] 'fast' is primarily attested as taking r [ɾ] (Swift 1963 , Wedel 1999, Kelepir 2001, Göksel & Kerslake 2006, Ketrez 2012, Özçelik 2012, Stachowski 2014, Demir 2017, van Schaaik 2020, Yılmaz 2020). However, some sources attest that it takes [p] (Wedel 1999, Özçelik 2012, Yılmaz 2020), [s] (Özçelik 2012, Yılmaz 2020), and least commonly, [m] (Taneri 1990). See below for a summary of combinations of possible linkers for words which have more than one attested linker.

(57) *Summary of linkers for words with multiple attested linkers*

Linkers	*# of TER forms*
p, s	35
m, p	25
m, p, s	9
m, s	7
r, s	5
p, r	2
p, r, s	2
m, p, r, s	2
m, r, s	1
p, s, m	1

In some cases, sources contradict each other about which linker should be chosen. For example, *top* [top] ‘round’ is reported as able to undergo TER in four sources, to gain the meaning ‘completely spherical’. Of those four, two report that it takes either [s] or *r* [ɾ] as a linker (Wedel 1999 and Özçelik 2012), one reports that it can take only [s] (Yılmaz 2020), and one reports that it can take only <r> [ɾ] (Stachowski 2014). All four of these sources report multiple linkers when multiple are attested. In other cases, however, the choice of linker seems very stable. For example, the *beyaz* [bejaz] ‘white’ is included in at least nine sources, and every source lists [m] as the only reported linker, including Taneri (1990), in which all 24 speakers who reduplicated the word chose [m] as the linker.

However, it is very possible that interspeaker variation is underreported in the dataset. Of the 113 total words with one attested linker, 39 (~35%) were attested in only one source. Of the 89 words with multiple linkers, on the other hand, only 4 (~5%) were attested in only one source. It is entirely possible then that many words which supposedly take only one linker, may in fact take more than one, if they were subject to further study.

2.5.13 TER conclusion

This section has laid out various past approaches to Turkish emphatic reduplication. These have included a fully lexicalized approach, in which the linker consonant is memorized with the base, as well as identity avoidance or dissimilation approaches. Additionally, various experimental data was laid out in which Turkish speakers reduplicated both real and nonce words. Finally, a review of reported data on TER linker consonants from a number of sources was provided. Overall, Turkish speakers show a large amount of variability in linker choice for both real and nonce Turkish words.

2.6 Plosive laryngeal alternation

Turkish shows alternation in laryngeal setting for plosives at morpheme boundaries. This section will lay out background and relevant past literature on the topic.

2.6.1 Laryngeal alternation in final plosives

Since this section will shift into a discussion specifically focused on consonants, I wanted to provide the Turkish consonant inventory for those less familiar with the language. Table 58 below shows the consonant inventory of Turkish, both using IPA and Turkish orthography. For simplicity, I am describing the laryngeal contrast in Turkish plosives in this section as 'voiced' and 'voiceless' and transcribing them as such. This is typically the convention used in other works describing the sound system of Turkish (see Lewis 1967, Underhill 1976, Clements & Sezer 1982, Inkelas & Orgun 1995, Stachowski 2014, Özçelik 2024, and others). In reality, the phonetic (and phonological) situation is likely more complex than this, which will be described in more detail in subsection 2.6.6 and in Chapter 6.

Note that /c ɟ/ have no consonant equivalent in Turkish orthography. When present phonologically (not due to allophony with [k g] adjacent to front vowels), these sounds can be distinguished with a hat on the vowel (Turkish: *şapkalı sesli harfler*). For example, *rüzgâr* /ɾyzcaɾ/ 'wind'. However, the hat on the vowel is frequently dropped in the orthography, leaving no differentiation between the velar and palatal stops in writing.

(58) *Turkish consonant inventory*

IPA	*Orthography*
Approximates: /l j/	Approximates: l y
Nasals: /m n/	Nasals: m n
Fricatives: /f v s z ʃ ʒ h/	Fricatives: f v s z ş j h
Tap: /ɾ/	Tap: r
Affricates: /ʧ ʤ/	**Affricates: ç c**
Stops: /p b t d k g c ɟ/	**Stops: p b t d k g**

The sounds of interest for this discussion are the plosives, that is, the affricates and stops (bolded). Turkish has a three-way pattern for plosives in final position, (59a-c):

(59) *Types of laryngeal setting in Turkish plosives*

	Phone	*Bare Form*	*Translation*	*Suffixed Form*	*Translation*
a.	p	küp [kyp]	'cube'	küpü [ky.p-ym]	'my cube'
	t	at [ɑt]	'horse'	atım [ɑ.t-ɯm]	'my horse'
	ç [tʃ]	saç [sɑtʃ]	'hair'	saçım [sɑ.tʃ-ɯm]	'my hair'
	k	bank [bɑnk]	'bench'	bankım [bɑn.k-ɯm]	'my bench'
b.	b	tab [tɑb]	'print'	tabım [tɑ.bɯm]	'my print'
	d	ad [ɑd]	'name'	adım [ɑ.d-ɯm]	'my name'
	c [dʒ]	sac [sɑdʒ]	'sheet metal'	sacım [sɑ.dʒ-ɯm]	'my sheet metal'
	g	ring [ɾing]	'boxing ring'	ringim [ɾin.g-im]	'my boxing ring'
c.	p ~ b	kap [kɑp]	'container'	kabım [kɑ.b-ɯm]	'my container'
	t ~ d	tat [tɑt]	'taste'	tadım [tɑ.d-ɯm]	'my taste'
	ç ~ c [tʃ ~ dʒ]	güç [gytʃ]	'strength'	gücüm [gy.dʒ-ym]	'my strength'
	k ~ g	renk [ɾenk]	'color'	rengim [ɾen.g-im]	'my color'

Some final plosives are always voiceless regardless of environment (examples in 59a). Other final plosives (59b) are always voiced. Finally, some final plosives (59c) alternate in laryngeal setting, being voiceless in the bare form and voiced when resyllabified to the onset of a vowel-initial suffix. The vast majority of final plosives are either always voiceless or alternating; always voiced final plosives are very rare in Turkish and are primarily loanwords.

If a root has a final voiceless plosive in the bare form, it is not possible to tell from the root alone whether the final plosive will be voiced or voiceless when resyllabified into an onset. This information must be memorized. And, in fact, Turkish dictionaries tend to note alternations in their entries for this reason. When there is an alternation from voiceless to voiced in a root-final consonant, dictionaries will mark the entry to indicate the alternation to the reader. See below for an example taken from an online dictionary entry.

(60) *Entry from Türk Dil Kurumu Sözlükleri showing alternation*

kanat, -dı

1. *isim* Kuşlarda ve böceklerde uçmayı sağlayan organ:
"*Bir daldan hızla geçip gitti bir kuş kanadı / Sessizlik bir tüy gibi bir dakika sallandı /*

No other Turkish consonants undergo this type of voicing alternation, a fact which will be crucial for the later analysis. For example, fricatives do not undergo any alternation with respect to laryngeal setting. A voiceless word-final fricative will *always* be voiceless, even when a vowel-initial suffix is added. In addition, voiced fricatives are not rare in final position as voiced stops are. Many Turkish words contain voiced fricatives in codas, e.g. *kez* [kez] 'occasion', *plaj* [plɑʒ] 'beach', and *ev* [ev] 'home'.

The distribution of alternating final roots consonants varies by place of articulation. In comparison to other plosives, there are relatively few examples of root final alternating alveolar plosives *t* ~ *d*. Göksel & Kerslake (2006) attribute this to the Arabic feminine suffix -et/-at having an unchanging voiceless [t], as well as the fact that the many t-final loanwords from French not alternating.

Notice that the examples in (59) with velar plosives are shown after [n]. This is because *k* ~ *g* alternation only occurs following [n]. Interestingly, *k* ~ *g* alternation does not occur following other sonorants; for example, all instances of word-final [lk] and [ɾk] seem to have a non-alternating voiceless [k] (e.g., *cılk* [dʒɯlk] 'rotten', *delk* [delk] 'friction', *halk* [hɑlk] 'folk'; *fark* [fɑɾk] 'difference', *terk* [teɾk] 'abandonment', *türk* [tyɾk] 'Turk'). When the only phoneme in a coda, velars alternate with what is written as *ğ* 'soft g' (Turkish: *yumuşak ge*) in the orthography (Göksel & Kerslake 2006). This process is labelled as *velar drop* (Inkelas & Orgun 1995). The phonological status of *ğ* and background on velar drop will be explained further in 2.6.2.

Alternation following sonorants is not exclusive to velars. Other plosives undergo alternation following sonorants as well:

(61) *Types of laryngeal setting in Turkish plosives following sonorants*

	Phone	*Bare Form*	*Translation*	*Suffixed Form*	*Translation*
a.	p	kulp [kulp]	'handle'	kulpum [kul.p-um]	'my handle'
	t	alt [ɑlt]	'bottom'	altım [ɑl.t-ɯm]	'my bottom'
b.	d	trend [tɾend]	'trend'	trendim [tɾen.d-im]	'my trend'
c.	p ~ b	kalp [kɑlp]	'heart'	kalbim [kɑl.b-ɯm]	'my heart'
	t ~ d	ant [ɑnt]	'oath'	andım [ɑn.d-ɯm]	'my oath'
	ç ~ c [tʃ ~ dʒ]	hınç [hɯntʃ]	'grudge'	hıncım [hɯn.dʒ-ɯm]	'my grudge'

In the table above, not every possible place of articulation is present for every type of laryngeal setting. The postalveolar affricates appear to only alternate before sonorants, as I was able to find a number of other examples which alternate (e.g., *felç* [feltʃ] 'stroke, paralysis', *malç* [maltʃ] 'mulch', *inç* [intʃ] 'inch', *tunç* [tuntʃ] 'bronze'), but none which are always *c* [dʒ] or always *ç* [tʃ] in this position. Examples with word-final always voiced consonants are quite rare, as mentioned earlier, and the number of those words with complex codas involving sonorants are even rarer. I was not able to find any with *b* or *c* [dʒ]. Aside from the one example in (28b), there appear to be a couple more involving the alveolar: *gard* [gaɾd] 'guard (sports)', and *bold* [bold] 'bold'. However, note that all the plosives *p t k ç* [p t k tʃ] can alternate in laryngeal setting following a sonorant consonant, so the environment conditioning the alternation cannot be accurately characterized as intervocalic.

In addition, it is important to note that a *following* sonorant consonant does not condition a change in alternating plosives. For example, when adding the sonorant-initial plural suffix to an alternating root, the final consonant will not change: *kaplar* [kɑp-laɾ] ‘containers’ (cf., **kablar* *[kɑb-laɾ]). This indicates that resyllabification is crucial for roots which undergo final plosive laryngeal alternation.

While the examples provided above are all nouns, there are also a limited number of verbal roots which undergo this alteration as well. To my knowledge, the alternating plosives in verbal roots are all alveolars. Below are the examples listed in Göksel & Kerslake (2006:16):

(62) *Examples of verbal roots with alternating t/d*

Bare Form	*Translation*	*Suffixed Form*	*Translation*
et- [et-]	do	ediyor [e.d-i.joɾ]	‘he/she/it is doing’
git- [git-]	go	gidiyor [gi.d-i.joɾ]	‘he/she/it is going’
güt- [gyt-]	herd	güdüyor [gy.-dy.joɾ]	‘he/she/it is herding’
tat- [tɑt-]	taste	tadıyor [tɑ.d-ɯ.joɾ]	‘he/she/it is tasting’

In addition, there are a number of compound verbs involving the auxiliary verb *et-* [et-] which also alternate, for example: *kaydet-* [kaydet-] ‘save’ becomes *kaydediyor* [kɑjdedijoɾ] ‘he/she/it is saving’, *bahset-* [bɑhset-] ‘to speak of’ becomes *bahsediyor* [bɑhsedijoɾ] ‘he/she/it is speaking of’.

Resyllabification is crucial in verbal forms as well. As with nouns, a following sonorant does not condition a change in laryngeal setting for the consonant of these verbs. This can be observed in the infinitival forms, since the infinitival suffix is sonorant-initial: *etmek* [et-mek] ‘to do’, *gitmek* [git-mek] ‘to go’, *gütmek* [gyt-mek] ‘to herd’, and *tatmak* [tɑt-mak] ‘to taste’.

Finally, there are some plosive-final suffixes in Turkish. These mostly consist of derivational suffixes, but there are a few examples of plosive-final inflectional suffixes as well. These plosives are all either non-alternating voiceless or alternating; I could not find any suffixes which ended in non-alternating voiced plosives. Place of articulation seems to have a large impact on the alternation. Suffixes ending in an alveolar never alternate and are never voiced (always *t*). Suffixes containing a postalveolar seems to always alternate (*ç* ~ *c* [tʃ ~ dʒ]). There are not, to my knowledge, any suffixes in Turkish ending in a bilabial plosive. Finally, there are a relatively large number of suffixes ending in *k* which alternate with *ğ*. The *k* ~ *ğ* alternation in suffixes will be set aside for now and discussed further in subsection 2.6.2. See below for examples of suffixes with alveolar and postalveolar final plosives, gathered from Göksel & Kerslake's (2006) list of Turkish suffixes:

(63) *Final suffix plosives*

Form	*Suffixed*	*Translation*	*Suffixed*	*Translation*
-at	tahkika**t** [takik-ɑ**t**]	'investigation'	tahkika**t**ı [takik-ɑ**t**-ɯ]	'investigation-ACC'
-It/-t/Ert	uyu**t**- [uju-t-]	'make/let [s.o.] sleep'	uyu**t**uyor [uju-t-ujoɾ]	'he/she/it is making/letting [s.o.] sleep'
-AC	süre**ç** [syɾ-e**tʃ**]	'process'	süre**c**i [syɾ-e**dʒ**-i]	'process-ACC'
-inC	basın**ç** [bɑs-ɯn**tʃ**]	'pressure'	basın**c**ı [bɑs-ɯn**dʒ**-ɯ]	'pressure-ACC'

2.6.2 Background on velar drop

As mentioned earlier, the velar plosive [k] patterns differently from other final plosives. When a [k] is the only phone in a final coda, it can alternate with *ğ*. However, [k] does not alternate with *ğ* in all cases. See below for examples of both alternating and non-alternating [k].

(64) *Examples of alternating and non-alternating [k]*

	Bare Form	*Translation*	*Suffixed Form*	*Translation*
a.	böcek [bœdʒek]	'insect'	böceğim [bœdʒe:im]	'my insect'
	çocuk [tʃodʒuk]	'child'	çocuğum [tʃodʒu:m]	'my child'
	gök [gœk]	'sky'	göğüm [gœ:ym]	'my sky'
b.	hukuk [hukuk]	'law'	hukukum [hukuk-um]	'my law'
	merak [meɾɑk]	'worry'	merakım [meɾɑk-ɯm]	'my worry'
	ok [ok]	'arrow'	okum [okum]	'my arrow'

There are also a number of suffixes with word-final *k* that alternates with *ğ*. One example is the diminutive -CIK (*Efecik* [efe-dʒik] 'Efe (diminutive)', *Efeciğim* [efe-dʒi:m] 'my dear Efe'[16]).

In addition, there are examples of final *g* ~ *ğ* alternation. For example, the *g* in *psikolog* [pisikolog] 'psychologist' alternates to *ğ* when resyllabified into an onset *psikoloğum* [pisikolo:um] 'my psychologist'. However, this alternation does not occur in all cases; some words have a non-alternating final *g* which does not change, e.g. *füg* [fyg] 'fugue' and *fügüm* [fygym] 'my fugue'.

There were two characters in Ottoman script which eventually became *ğ* in modern Turkish orthography, <غ> and <ڭ> (Ünal-Logacev et al, 2019). The former was used following back vowels and was pronounced as a voiced velar fricative [ɣ]. The latter was used following front vowels and was pronounced as a palatal glide [j]. These were combined into one character *ğ* during the reform of the Turkish writing system in the 1920s (see Özçelik 2024 Chapter 1). In native Turkish words, *ğ* and *g* are in complementary distribution: *g* is found word-initially and in onsets following a consonant, while *ğ* is found elsewhere (Ünal-Logacev et al. 2019, Özçelik 2024:35). However, the two sounds are no longer found

[16] This possessive construction is commonly used when talking directly to the person being referred to in the diminutive.

in complementary distribution due to loanwords with intervocalic and final *g* (e.g., *füg* [fyg] 'fugue', *Ege* [ege] 'Aegean Sea').

In some dialects of Turkish, the pronunciation of *ğ* is still [ɣ] (Inkelas & Orgun 1995:767, Göksel & Kerslake 2006:7). However, the realization of *ğ* is debated in modern İstanbul Turkish. Ünal Logacev et al. (2019) conducted a study of 24 native İstanbul Turkish speakers, 16 of whom were young and 8 of whom were elderly. Participants were recorded producing *ğ* in different positions within a word and in different vowel environments. No evidence of a velar gesture was found; for both age groups, the only phonetic correlate of *ğ* was a lengthening of the previous vowel.

Another study of 14 native Turkish speakers did find some evidence of a velar glide [ɰ] as measured by F1 and F2 values (Uzun 2021). However, the goal of this study was to examine glides in Turkish broadly. Because of this, results of this study are a bit more difficult to interpret with respect to *ğ*. First, the study did not control for dialect of the participants. Secondly, because the goal was to examine glides in general and not just *ğ*, the word list included some words which were presumed to contain a velar glide [ɰ] but did not include *ğ*, like *kauçuk* 'rubber'.

Phonologically, *ğ* seems to retain its consonant status in modern İstanbul Turkish, at least on some level. Ünal Logacev et al. (2019) bring up a good example demonstrating this. In Turkish, some otherwise vowel-initial suffixes will surface with a consonant when affixed to a vowel-final word. For example, the dative surfaces as just a vowel when affixed to a consonant-final word (e.g., *okula* [okul-a] 'to school'), and as a vowel following a palatal glide when affixed to a vowel-final word (e.g., *mağazaya* [ma:za-ja] 'to the store'). When a word ending in *ğ* is suffixed with one of this set of suffixes, no consonant is added (e.g., *dağa* [da:] 'to the mountain', cf **dağya* *[da:-ja]).

This dissertation will not include a phonological analysis of velar drop. Rather, the analysis given in Chapter 6 will cover alternation in the laryngeal setting of plosives only. Velar drop, while seemingly a related or similar phenomenon, will be left as an area for future research.

2.6.3 Laryngeal alternation in initial suffix onsets

In addition to the final plosives outlined above, there is a similar three-way pattern for plosive-initial suffixes. Unlike the final plosives, no resyllabification is taking place for these suffix initial plosives; they are always in onset position. Alternating plosive-initial suffixes always match the laryngeal setting of the preceding sound.

(65) *Alternating Plosive Initial Suffixes*

a.	-DA	evde [ev.-de] ‘at home’	sokakta [so.kɑk.-tɑ] ‘in the street’
b.	-CA	çocukça [tʃo.dʒuk.-tʃɑ] ‘childish’	aptalca [ɑp.tɑl.-dʒɑ] ‘idiotic’
c.	-GAn	etken [et.-ken] ‘factor’	sürüngen [sy.ɾyn.-gen] ‘reptile’

To my knowledge, there are no suffixes in Turkish with alternating *p* ~ *b* in initial position. Note that in (65c), *k* alternates with *g* and does not alternate with *ğ*. This is because *ğ* does not occur in onset position following a consonant. There are also plosive initial suffixes which do not alternate, both as always voiced and always voiceless.

(66) *Unchanging Plosive Initial Suffixes*

Form	*Suffixed (after voiced)*	*Translation*	*Suffixed (after voiceless)*	*Translation*
-baz	ateş**b**az [ɑteʃ-**b**ɑz]	'fire performer'	düzen**b**az [dyzen-**b**ɑz]	'cheat'
-tAy	Yargı**t**ay [jaɾgɯ-**t**aj]	'Supreme Court'	çalış**t**ay [tʃɑlɯʃ-**t**ɑj]	'workshop'
-gil	turunç**g**iller [tuɾuntʃ-**g**il-leɾ]	'citrus fruits'	bakla**g**iller [bɑklɑ-**g**il-leɾ]	'legumes'
-ken	okur**k**en [okuɾ-**k**en]	'while reading'	çocuk**k**en [tʃodʒu**k**:en]	'during childhood'

To my knowledge, there are no suffixes with unchanging initial *p*, *c* [dʒ], or *ç* [tʃ]. There are a couple of suffixes which potentially have a suffix-initial unchanging *d*: *-dan* and *-dar*, both suffixes of Persian origin (Göksel & Kerslake 2006:59). However, I was only able to find instances of these suffixes following voiced sounds, so it is not possible to discern if they would surface as *t* following a voiceless sound. Overall, there exists in plosive initial position a three-way pattern where some plosives are always voiceless, some are always voiced, and some alternate in laryngeal setting.

2.6.4 Past approaches

There are two main approaches to representing this alternation phonologically. The first is an underspecification approach, and the second is a coda devoicing approach. There is a third approach taken by Avery (1996), which is similar to, but distinct from, the devoicing approach. I will summarize all three approaches here.

This first approach makes use of underspecification for alternating plosives. Petrova et al. (2006), Kallestinova (2004), and Inkelas & Orgun (1995) analyze always voiceless final plosives as underlyingly voiceless, always voiced final plosives as underlyingly voiced, and final plosives which alternate as underlyingly underspecified. Alternating suffix plosives (as in the locative -DA) are analyzed in various ways: as underlyingly the same as root-final alternating plosives; that is, underspecified for voicing (Kallstinova 2004), as underlyingly voiced (Petrova et al. 2006), or simply not mentioned (Inkelas & Orgun 1995).

The second kind of analysis, found in several grammars, describes Turkish as having coda devoicing or disallowing voiced plosives in codas (Kornfilt 1997:491, Underhill 1976:41, Göksel & Kerslake 2006:15). In these accounts, always voiceless final plosives are underlyingly voiceless. Always voiced and alternating final plosives are both underlyingly voiced in this approach, with the non-alternating voiced final plosives being an exception to coda devoicing. The alternating suffix initial plosives are described in these grammars as matching the laryngeal setting of the preceding sound. However, underlying phonological representations are not given in these sources.

Özçelik (2024) also argues for the coda devoicing approach; however, Özçelik (2024) deals with the exceptionally voiced coda plosives in a unique way. Özçelik (2024) observes that many of the words with voiced final plosives would have homophones if they were devoiced (e.g. *ad* [ɑd] 'name' and *at* [ɑt] 'horse'). Özçelik (2024) then argues for a high ranking *Homophony constraint, which assigns a violation if coda devoicing would result in a homophone. This constraint would block coda devoicing for the majority of words with voiced final plosives, especially the portion that are monosyllabic.

Avery (1996) takes an approach which is similar to the coda devoicing approach. Avery does argue for a rule which blocks <b d g c> occurring in codas. However, Avery takes the position that [voice] is not the contrastive feature for Turkish plosives. Rather, Avery argues that the contrastive feature for plosives is [spread], and thus that a coda fortition rule is appropriate. To account for words with <b d g c> in codas, Avery proposes that these stops have Sonorant Voice; essentially, in their laryngeal setting they are grouped with sonorants, and so the coda fortition rule (which applies to plosives) will not apply to them.

2.6.5 What is going on with the Turkish laryngeal system?

Prior to proceeding with analysis, there is a deeper question which must be addressed about Turkish: what is the phonologically active laryngeal dimension for Turkish consonants? Phonetically, Turkish has been described as having prevoicing on <b d g> and aspiration on <p t k> (Öğüt et al. 2006; Cho, Docherty, & Whalen 2019; Göksel & Kerslake 2006). Phonologically, Turkish has been described as having a voicing contrast for plosives (Cho, Whalen, & Docherty 2019, Inkelas & Orgun 1995), and as having an aspiration contrast (Avery 1996), and finally as both a voicing and an aspiration contrast (Petrova et al. 2006, Kallestinova 2004, Nicolae & Nevins 2016).

If Turkish is indeed a language with a voicing contrast for plosives, it would be the only known exception to the generalization that all languages with a contrastive voice feature have regressive assimilation (see van Rooy & Wissing 2001). To my knowledge, there is no regressive voicing assimilation in Turkish. There are very few suffixes with (assuming a voicing contrast) always voiced initial plosives and fricatives <b d g v z>, but those that are present do not seem to condition voicing on a preceding consonant. The suffixes -*vari*, -*zade*, and -*gil* do not seem to cause voicing assimilation of a previous plosive.

2.6.6 General phonetic studies of Turkish stops

There have been a number of phonetic studies conducted on Turkish plosives, focusing specifically on stops.

Kallestinova (2004) collected recordings from one native Turkish speaker, and recordings from a second Turkish speaker were taken from the University of Victoria Phonetic Database (UVPD). Words were produced in isolation. Kallestinova (2004) then examines VOT of stops in different positions. For both speakers, word-initial <b> was almost always prevoiced, while word-initial <d g> were almost always voiceless. Data from the one native Turkish speaker was collected to look at intervocalic stops (the data from UVPD did not include enough intervocalic stops). In contrast to the word-initial stops, intervocalic <b d g> were always voiced during closure by that speaker, with most tokens being voiced through 100% of the closure. In both the word-medial and word-initial data, <p t k> had long-lag VOT (30+ msec). Kallestinova (2004) also includes a more targeted investigation of plosives which alternate in laryngeal setting; we will return to this study and summarize those results in the next section.

Öğut et al. (2006) further investigated VOT of Turkish word-initial stops. In the study, 30 native Turkish speakers produced monosyllabic stop-inital words in a sentence frame[17]. See below for a summary of the average VOT values across all speakers.

(67) *Summary of results, adapted from Öğüt et al. (2006:1097, Table 2)*

Orthography	*VOT (msec)*	*Stan. Dev. (msec)*
p	40.7	17.6
t	49.7	10.3
k	69.4	10.8
b	-65.9	39.7
d	-52.9	41.6
g	-9.7	22.9

While mean VOT values for <b d g> are all negative, the mean VOT for <g> is barely negative. In fact, Öğüt et al. (2006) noted that 40% of <g> tokens were pronounced with no prevoicing at all.

[17] The sentence frame was not reported in the study, so it is unclear what environment the stops were pronounced in (intervocalic, after a consonant, etc).

Kılıç (2018) also collected data on VOT of Turkish stops, with a different purpose: to examine how VOT is modulated across languages for native speakers of Turkish who learned English as a second language. In most dialects of English, <b d g> are not commonly prevoiced, being pronounced as voiceless plain (Lisker & Abramson 1967, Neary & Rochet 1994). Sixteen native Turkish speakers who spoke English at an advanced level and sixteen native English speakers participated in the study. Native Turkish speakers were recorded pronouncing stop-initial words in both Turkish and English. Native English speakers pronounced stop-initial words in English. All words were spoken in isolation, without a sentence frame. The table below shows a summary of the VOT results.

(68) *Summary of VOT results, adapted from Kılıç (2018:278)*

Orthography	*Native Turkish VOT (msec)*	*Non-native English VOT (msec)*	*Native English VOT (msec)*
p	40.3	45.7	45.7
t	50.0	52.7	65.6
k	70.5	72.2	69.6
b	-63.4	-46.2	11.4
d	-50.9	-31.8	16.8
g	-11.0	-3.3	27.2

The smaller table below summarizes the change in VOT for native Turkish speakers when producing stops in Turkish as compared to producing the same stop in English.

(69) *Amount of VOT change from native Turkish to non-native English*

Orthography	*VOT change (msec)*
p	+5.7
t	+2.7
k	+2.3
b	+17.2
d	+19.1
g	+7.7

One interesting finding of the study is that, as in Öğüt et al. (2006), mean VOT for <g> for native Turkish, while negative, is quite close to zero. Another result of note is that the negative VOT on <b d g> differs between native Turkish and non-native English. For Turkish speakers, there is clearly a decrease in prevoicing when speaking non-native English as compared to native Turkish. This indicates that native Turkish speakers are able to modulate their degree of prevoicing when speaking English. For <p t k>, native Turkish, non-native English, and native English all have comparable VOT values, all being pronounced as aspirated with long-lag VOT.

Ünal-Logacev et al. (2018) examined production of Turkish plosives through both acoustic and articulatory measures. The study examined the production of plosives <t d ç c> by six native Turkish speakers. The phones of interest were produced in a sentence frame and in word initial position between two vowels. Acoustic measures included consonant duration, closure duration, voicing duration, percentage of voicing into closure, and VOT (for the stops only). Articulatory measures were collected with an electropalatography (EPG) plate inserted into the mouth. The EPG measured intra-oral pressure and tongue-palatal contact.

Results showed that <d> and <c> were voiced through the entirety of closure in basically all instances, while <t> and <ç> were voiced for about 25% of closure on average. As expected, <t> had a positive VOT, while <d> had a negative VOT. In addition, <d c> had a shorter duration than that of <t ç> on average. In terms of non-accoustic measures, <t ç> had a much higher maximum velocity of intraoral pressure. No difference was found for tongue palatal contact measures.

Finally, Feizollahi (2009) examined Turkish laryngeal stop contrasts through both production and perception. For the production portion, Feizollahi (2009) examined the production of Turkish stops in varying contexts by four Turkish speakers. Results showed that word-initial <b d g> were voiced following a voiced consonant and not voiced following a voiceless consonant. Word-initial <p t k> were voiceless and aspirated in both contexts. When following a voiceless consonant, the mean VOT difference between <b d g> and <p t k> respectively was quite small, varying from 6-16 msec depending on PoA and speaker.

Due to this small VOT difference, Feizollahi (2009) conducted a perception study to investigate whether other cues aside from VOT may be used by speakers to maintain the laryngeal contrast in stops following a voiceless consonant. The perception study stimuli were created using software to manipulate the VOT of word-initial <p t k b d g> following a voiceless stop. Results showed that speakers were able to reliably distinguish laryngeal category in unaltered stops, demonstrating that the contrast was maintained, despite the small difference in VOT. For altered files, the VOT crossover point for <b d g> was significantly higher than that of <p t k> (for instance, an altered <b> was perceived as a <p> at a VOT of ~15 msec, while an altered <p> was generally distinguished as a <p> even when VOT was 0 msec). These results suggest that Turkish speakers are indeed using other (non-VOT) cues to distinguish the laryngeal category of a stop.

2.6.7 Targeted phonetics of Turkish plosive alternation

There have been several studies which have specifically investigated the phonetics of Turkish with the goal of better understanding plosive laryngeal alternation.

Kallestinova (2004), discussed earlier in the previous subsection, was also interested in whether the voiceless allophone of the alternating stops have any detectable phonetic difference from the always voiced final plosives. To that end, Kallestinova compared the VOT of alternating word-final stops to always voiceless final stops for one speaker (e.g., the always voiceless t in *at* [ɑt] 'horse' versus the alternating t in *kanat* [kɑnɑt]). Results showed no difference between the VOT of the two kinds of stops, they were all found to have log-lag VOT.

Finally, Kallestinova (2004) compares the voiceless allophones of alternating suffix stops to those of always voiceless suffix stops for one speaker. The VOT for the alternating is in general shorter than the always voiceless suffix stops. Kallestinova then concludes that alternating suffix stops must be plain, while non-alternating voiceless stops are aspirated. However, there is a notable confound in this study: all of the alternating suffix stops were <t>, while all of the non-alternating suffix stops were <k>. There is a well known place of articulation effect on VOT, where generally stops produced with an obstruction farther back in the vocal tract (like velars) are known to have a longer VOT than those produced farther forward (like alveolars) (Lisker & Abramson 1967, Neary & Rochet 1994). This effect has also been documented in Turkish specifically for word-initial stops (see the previous discussion of Öğüt et al. 2006, Kılıç 2018). As a result, it is impossible to tell from Kallestinova's results whether the difference in VOT between the types of suffix stops is due to a difference in underlying representation, or merely a phonetic difference based on place of articulation.

Wilson (2003) investigated the laryngeal alternation in Turkish stops by recording three native speakers producing words with alternating, always voiced, and always voiceless final stop consonants. The first speaker, whose productions were recorded and analyzed before the other two speakers, showed no differences (in terms of duration and voicing into closure) between alternating stops and always voiceless stops in the bare form, indicating that neutralization is complete. This is consistent with the aforementioned finding from Kallestinova (2004) that alternating and always voiceless final stops were not different in terms of VOT. The always voiced stops, however, differed from the other two categories in terms of voicing into closure. Wilson (2003) further reported that the first speaker showed a difference between always voiced and alternating stops in intervocalic position, where the alternating stops were longer in duration than always voiced stops. The two other speakers were recorded in an effort to replicate the latter result, but no duration differences between alternating and always voiced stops in intervocalic position were found for either of the additional speakers.

Finally, Nicolae & Nevins (2016) conducted two experiments involving the production of Turkish fricatives. In the first study, eight native Turkish speakers were asked to produce words ending in voiced and voiceless fricatives [f v s z ʃ ʒ]. Results indicated that fricatives retained their laryngeal contrast in word final position (unlike stops, as observed by Kallestinova 2004 and Wilson 2003).

For the second study, Turkish and Russian native speakers were recorded producing stops [t d] and a fricative [s] immediately preceding sonorant consonants [l n]. For Turkish speakers, the voiceless fricative [s] induced more devoicing on a following sonorant than did [t d]. In addition, Turkish speakers devoiced the following sonorant much more following [s] than did Russian speakers (at least 65% devoicing compared to at most 15% for Russian speakers). From this result, Nicolae & Nevins (2016) argue that Turkish obstruents are overmarked. They propose that fricatives are marked as [voice] or [spread glottis]. For plosives, they argue that Turkish has a three-way contrast between [voice] (always voiced final plosives), [spread glottis] (always voiceless final plosives), and unspecified for laryngeal setting (alternating final plosive). They argue that this difference between the fricatives and plosives (fricatives having a two-way contrast and plosives having a three-way contrast) is why plosives alternate but fricatives do not.

2.6.8 Cross-linguistic comparisons

Cho, Whalen, & Docherty (2019) conducted a cross-linguistic study of dento-alveolar stops across 19 languages, one of which was Turkish. While Turkish was characterized as a voicing language, due to its prevoicing on <d>, it was also noted that Turkish 'voiceless plain' <t> was a 'rightward deviation from the general pattern' of voiceless plain stops in their study (Cho et al. 2019:55). Turkish <t> had an average VOT of 41ms, double the next highest average VOT value for a voiceless plain [t] (21ms in American English)

Phonetically, Turkish seems somewhat similar to Swedish. Swedish, like Turkish, has a two-way laryngeal contrast in stops. In word initial position, Swedish stops have been characterized as having long-lag VOT on <p t k> and prevoicing on <b d g> (Helgasen & Ringen 2008, Ringen & Suomi 2012).

2.6.9 Plosive laryngeal alteration conclusion

This section has provided background on the phenomenon of laryngeal alternation in Turkish plosives, along with a brief summary of previous phonological literature on the topic. In addition, further context about past research into the phonetics of Turkish plosives was given. In general, results of these studies indicate that Turkish <b d g> are often prevoiced, though this varies based on place of articulation and on context. In addition, based on limited data from two studies (Wilson 2003 and Kallestinova 2004), it seems that in word-final position, alternating <p ~ b> and <t ~ d> are produced identically to always voiceless <p>

and <t>, respectively. These facts will be revisited in Chapter 6, where an analysis of plosive laryngeal alternation will be provided.

2.7 Literature review conclusion

This section has provided background information and relevant literature related to the main topics covered in this dissertation: vowel harmony, vowel alternation in coda clusters, emphatic reduplication, and plosive laryngeal alternation. The following four chapters will provide an analysis of each of these using the EFP framework.

3 Vowel Harmony

The goal of this chapter is to provide a comprehensive view of Turkish vowel harmony (TVH) which includes 'exceptions' to TVH. This chapter outlines the EFP graph representations for harmonic words, as well as both kinds of exceptions introduced in the literature review: invariable vowels (in both roots and suffixes), and exceptional front or back suffixes. Finally, some related topics to TVH, prefixes and supposedly epenthetic vowels, will be discussed as well.

3.1 Representing the harmonic case

This work makes use of features from Avery & Idsardi (2001) which are underspecified and privative (see section 2.2 of the literature review for further details). The three contrastive vowel features of Turkish (height, frontness, and rounding), can be represented using this feature system in the following way using dimensions and completion rules:

(70) *TVH completion rules*

	Dimension		*Gesture*
a.	Tongue Thrust (TT)	→	[front]
b.	Tongue Height (TH)	→	[high]
c.	Labial (LAB)	→	[round]

The absence of a dimension is represented by a bare dimension node, called a superordinate marking (see Raimy 2021), indicative of the fact that the articulator is in a neutral state:

(71) *TVH superordinate markings*

	Superordinate marking		*Articulatory Consequence*
a.	DORSAL^{TT} (DOR^{TT})	→	non-front (central or back)
b.	DORSALTH (DORTH)	→	non-high (mid or low)
c.	ORAL PLACE^{LAB} (OP^{LAB})	→	non-round

As an example, a vowel /i/ could be represented phonologically with the following dimension specifications: TT, TH, and OP^{LAB}. During pronunciation, the completion rules will be used, and the vowel will be pronounced as front, high, and non-rounded. See below for a table of all the Turkish vowels, organized by their dimensions and superordinate markings.

(72) *Turkish vowels with dimensions & superordinate markings*

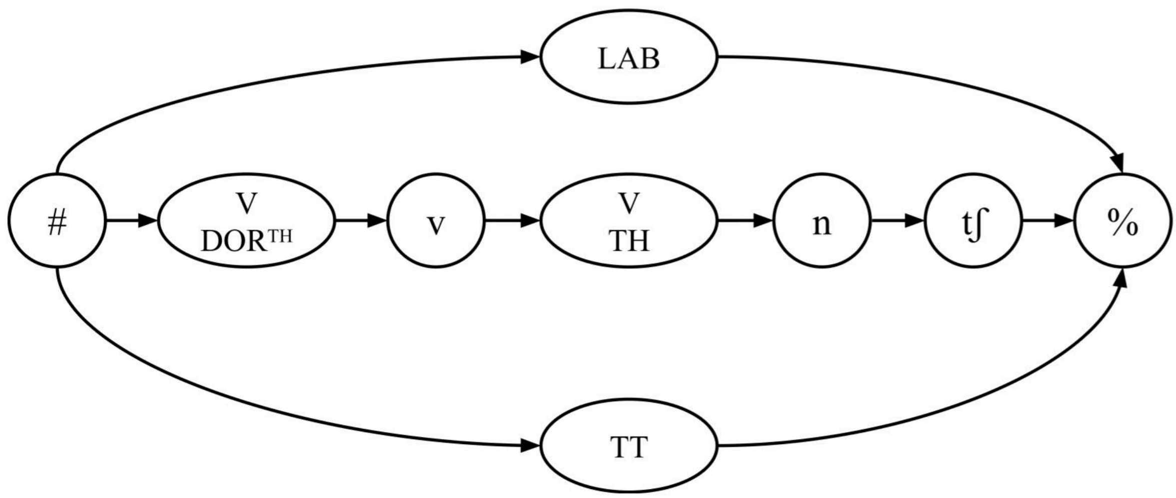
	Tongue Thrust		Dorsal^{TT}	
	Labial	Oral Place^{LAB}	Labial	Oral Place^{LAB}
Tongue Height	*ü* /y/	/i/	/u/	*ı* /ɯ/
DorsalTH	*ö* /œ/	/e/	/o/	*a* /ɑ/

The actual representations of Turkish vowels will end up being a bit more nuanced than this, because I will make use of autosegmental phonology- but let this suffice as an example to illustrate dimensions and gestures.

As a reminder, TVH operates on two of the three dimensions: TT (frontness) and LAB (rounding). TH (height) is entirely lexical and not a result of any harmony processes. However, the TH dimension does interact with the harmony system in that only TH (high) vowels are subject to rounding harmony, while DORTH (non-high) vowels are not.

Phonological representations of morphemes which fully conform to TVH can be constructed using the EFP framework (see section 2.1 of the literature review for further details on EFP directed graphs). Below is a graph representing the harmonic Turkish root *ovünç* [œvyntʃ] 'pride'.

(73) *EFP graph of övünç [œvyntʃ] 'pride' and its precedence relationships*



Rounding autosegmental stream (top)

precedes LAB

LAB precedes %

Frontness autosegmental Stream (bottom)

precedes TT

TT precedes %

Deautosegmentsalized stream (middle)

precedes /vowel, DORTH/

/vowel, DORTH/ precedes v

v precedes /vowel, TH/

/vowel, TH/ precedes n

n precedes tʃ

tʃ precedes %

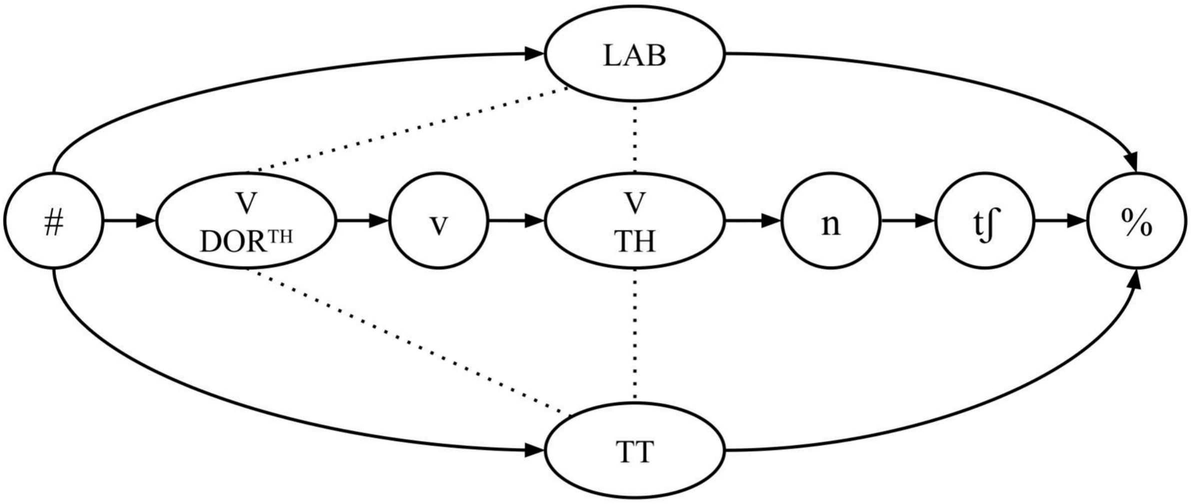
output: *övünç* [œvyntʃ]

There are separate streams in the graph above to represent the autosegmental tiers that carry harmony information. The LAB (rounding) harmony stream is on the top and the TT (frontness) harmony stream is on the bottom. Since all vowels in this word match for frontness and rounding, there is only one event in each of the autosegmental streams. Finally, the middle stream holds the de-autosegmentalized information. Because TH (height) is lexical and not a product of harmony processes, it must be de-autosegmentalized and placed inside the vowel events.

An important representational point here is that, in Turkish, the de-autosegmentalized part of a vowel consists of only a vowel specification and height information. All harmony information (rounding and frontness) is encoded in parallel streams. This is fundamentally very similar to the representations of Turkish vowels used by Clements & Sezer (1982), who posit an underlyingly high vowel *I* and an underlyingly non-high vowel *A*, with rounding and frontness represented in autosegmental tiers.

The key to understanding the EFP representation is that during pronunciation, the underspecified vowel phonemes receive their missing features (LAB and TT) from the streams that scope over them. So, when the example above is pronounced, /vowel, DOR^{TH}/ and /vowel, TH/ will seek out their missing frontness and rounding information from the autosegmental streams, causing them to be filled in with TT and LAB, yielding *ö* [œ] and *ü* [y], respectively. This is similar to the account of vowel harmony laid out by Nevins (2010) which was mentioned in the literature review. In Nevin's (2010), underspecified harmonizing vowels 'seek out' rounding and frontness information. However, in this account, rather than seeking out the information from the nearest vowel, the information is sought out from the harmony streams (for frontness and rounding) which scope over the vowel. The graph in (74) below shows the example in (73) dotted lines between the vowels and autosegmental harmony events (LAB and TT) which scope over them. Note the dotted lines here are for illustrative purposes only, to make it intuitive for the reader to see which events scope over the vowels; I am not positing some kind of special 'dotted line edge' in the graph.

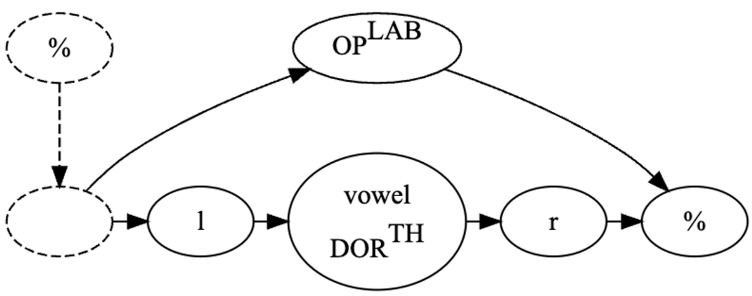
(74) *övünç [œvyntʃ] 'pride' with dotted lines connecting vowels and harmony events*



It is worth noting here that the choice to represent vowel harmony as being active in some way within roots is not uncontroversial; in fact, whether or not TVH remains active within roots is debated (see Kabak 2011 as well as section 2.3.6 of the literature review for a discussion of this debate). We will return to this fact later, in our discussion of exceptions.

Next, EFP graph structure can be used to represent TVH in suffixation. Below is an EFP directed graph representing the plural suffix *-lAr* /-lAɾ/. This suffix has a non-high harmonizing vowel which will harmonize for frontness of the preceding vowel, but is invariably unrounded.

(76) -lAr */-lAɾ/ plural suffix*



The dashed line parts of the graph represent the part of the morpheme that matches or slots on to a particular lexical item. In this case, the dashed part is the 'sticky end' of the suffix which attaches to the base. Here, the sticky end matches to the sink (%) so the suffix will attach to the end of a base, as expected. The affix will merge with the sink attached to the base, so the sink will only appear once in the combined stem + suffix form. Note that this is different from the way Papillon (2020) represented suffixation in PROP (see subsection 2.1.3 of the literature review for an overview of Papillon's approach to suffixation).

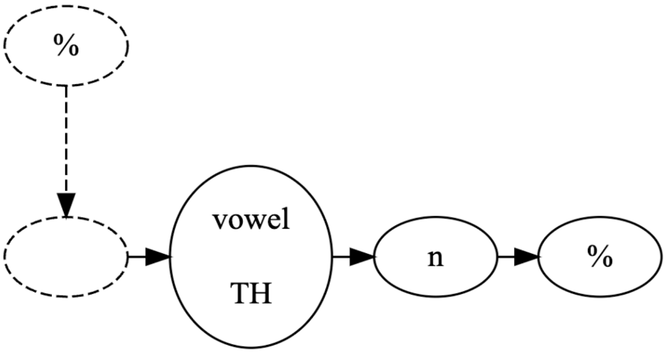
This suffix has an OP^{LAB} event in the autosegmental stream for rounding harmony. Recall that OP^{LAB} is a bare Oral Place dimension node- meaning that the lips will not undergo rounding, but rather say in a neutral state. This stops a LAB event from the stem (if present) from being sequenced during the suffix, meaning that the suffix vowel will not surface as round. This is necessary because non-high harmonizing vowels (like the A in *-lAr* /-lAɾ/) are always unrounded, even when suffixed on a stem with round vowels.

Lastly, note that the suffix has no autosegmental stream for TT (frontness), nor is the TT dimension contained within the vowel event: it is completely lacking from the graph. This is because suffixes are entirely dependent on the preceding vowel for determining frontness[18]. This omission represents the fact that *-lAr* /-lAɾ/ is 'hungry' for a frontness stream (this concept is also very akin to what is laid out in Nevins 2010).

There is another kind of harmonizing suffix in Turkish which are high vowel harmonizing suffixes, which undergo harmony for both frontness and founding. Shown below in (77) is an EFP graph of a suffix with a high harmonizing vowel: the second person singular possessive suffix *-In*.

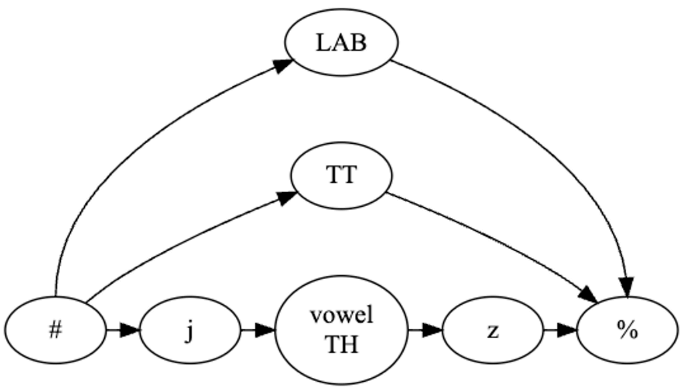
[18] This is not the case for exceptional non-harmonizing suffix vowels, which are invariable (e.g., the second vowel of the present progressive *-Iyor* /-Ijoɾ/, which is always /o/). We will discuss those cases in the following sections.

(77) -In *2nd sing possessive suffix*



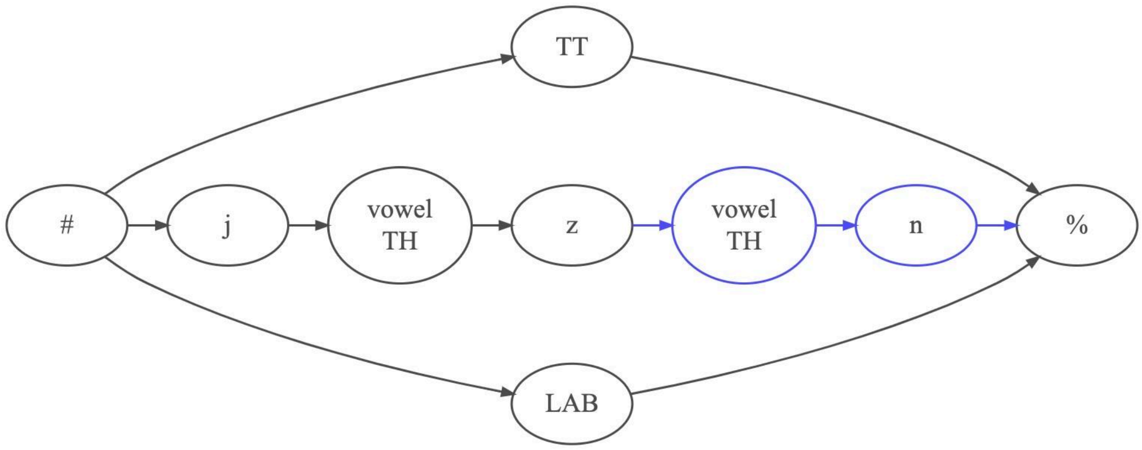
The only dimension which is specified for the vowel is TH, indicating that it is high. There are no autosegmental streams for the LAB or TT dimensions. Phonologically, this omission encodes that high harmonizing suffixes match the preceding vowel for both rounding and frontness. The vowel is 'hungry' for both LAB (rounding) and TT (frontness) streams and will gain that information from the preceding vowel. Now that both suffixes have been introduced, graph structures will be shown to demonstrate what occurs when these suffixes are attached to a base. Below is the example bare form *yüz* [jyz] 'face' which both of the suffixes given above will be attached to.

(78) *EFP graph of yüz [jyz] 'face'*



As with the earlier example *övünç* [œvyntʃ], *yüz* [jyz] has TT and LAB events in their respective autosegmental streams which encode the frontness and rounding of the vowel. So far, this is not very interesting, since *yüz* [jyz] has only one vowel. The bare form can be suffixed with the suffixes given above. The below graph shows the second person singular suffix -In affixed *yüz* [jyz] to form the multimorphemic word *yüzün* [jyz-yn]:

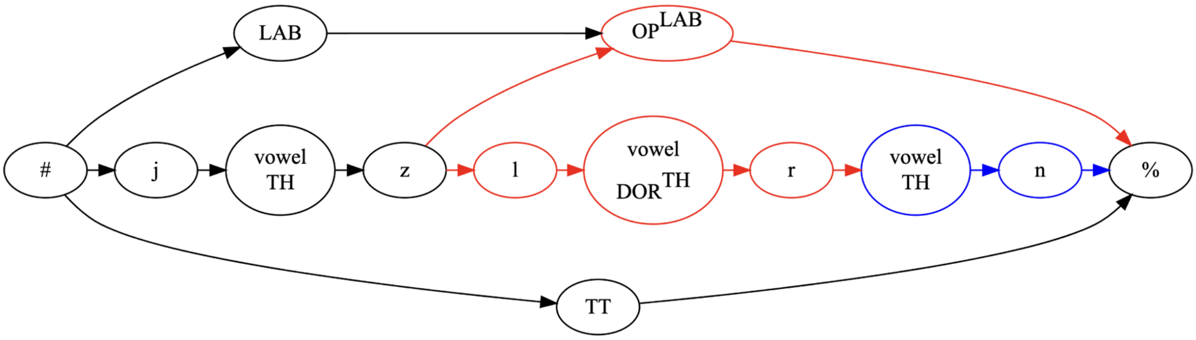
(79) *yüzün [jyz-yn] 'your (sing) face'*



Since the suffix has no frontness or rounding specification of its own, the TT and LAB events scope over the high harmonizing vowel in the second person singular (blue). This causes the vowel to be pronounced as *ü* [y], harmonizing for both frontness and rounding with the root, as expected for a high harmonizing vowel.

Next, the plural suffix *-lAr* /-lAɾ/ can be added to form the word *yüzlerin* [jyz-leɾ-in] 'your (sg) faces'. Below in (80) is the EFP graph structure for this form:

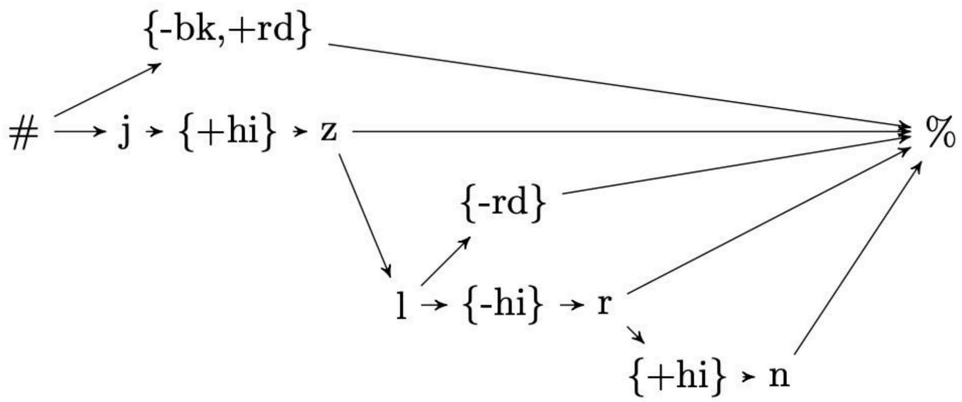
(80) *yüzlerin [jyz-leɾ-in] 'your (sg) faces'*



The OP^LAB event attached to the plural suffix *-lAr* /-lAɾ/ (red) scopes over the following second person singular suffix *-In* (blue). In this way, the OP^LAB event attached to the plural suffix blocks the spread of LAB (round) onto the second suffix. This representation has successfully encoded rounding opacity- where a low harmonizing vowel (as in the plural suffix *-lAr* /-lAɾ/) blocks rounding harmony from the base spreading to following high harmonizing vowels.

This is one place where the representation of harmony diverges from Papillon (2020). As mentioned in the literature review (section 2.1.3), Papillon (2020) makes use of nested autosegmental streams in his account of TVH.

(81) *Representation of yüzlerin [jyz-leɾ-in] from Papillon (2020:28, example 22)*



In the above graph, the {-round} autosegment (equivalent to OP^{LAB} in the graph in 80) has no direct precedence relationship with {+round} (equivalent to LAB). The {-round} autosegment is nested under the {+round} autosegment which scopes over the entire word (follows the source # and precedes the sink %). In Papillon (2020)'s account, a vowel will receive autosegmental features from the innermost nested autosegmental stream. So, the second vowel in the word above will be {-round}, since the {-round} autosegmental stream is the innermost for that vowel.

In this dissertation, a precedence relationship between the LAB and OP^{LAB} autosegments is added. In addition, the earlier autosegment, LAB does not scope over the whole word, but rather just over the vowels in the word which it applies to. This choice was made because Turkish has an opaque vowel harmony system. Once LAB is blocked from spreading by a OP^{LAB} event, it will not resurface again unless a vowel which is specified for LAB is present later in the word (for example, if the present progressive suffix -Iyor [Ijoɾ], with an invariably rounded vowel, is added). So, there is no instance where the outermost autosegmental feature would resurface, and thus no reason for the feature to scope over the entire word.

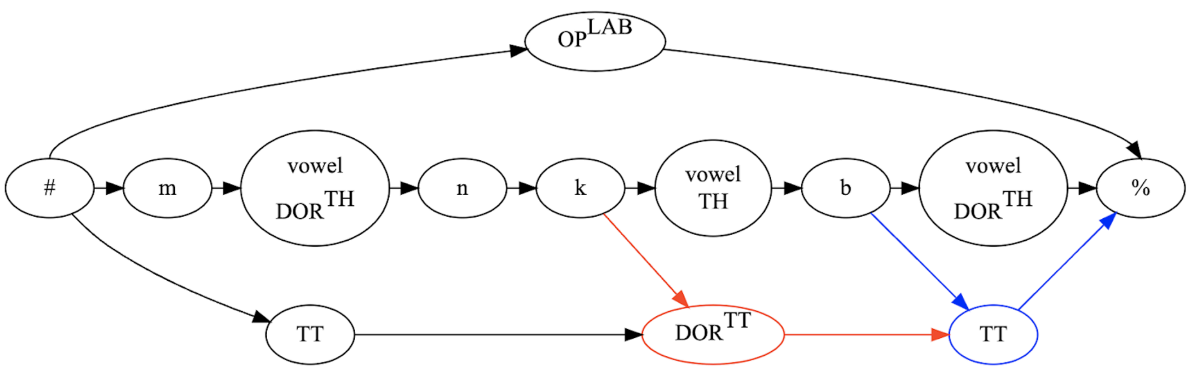
3.2 A comprehensive account of Turkish vowel harmony

Thus far, only EFP representations for morphemes which fully conform to TVH have been shown. The EFP system, while unconventional, has not done anything particularly novel or exciting yet. In fact, many of the key points that we have gone through so far were essentially laid out in Papillon (2020) (with some minor differences as described above). However, the examples we have seen so far do not show the full story, as there are numerous so-called exceptions to TVH. In this section, I lay out how the EFP graph representation can be used to encode so-called exceptions to TVH in the phonology.

3.2.1 Representing exceptional root words & immutable suffix vowels

Invariable disharmonic vowels can be represented as follows. For every vowel which does not conform to TVH, project from the segment immediately preceding it a new dimension setting into the appropriate autosegmental stream. All disharmonic vowels can be encoded in this manner, including those involving *ı ü ö* /ɯ y œ/, which Clements & Sezer (1982) discount. For example, consider the EFP graph representation of the word *menkıbe* [menkɯbe] 'legend' in (82) below, which is an 'exception' to TVH, having a back vowel between two front vowels:

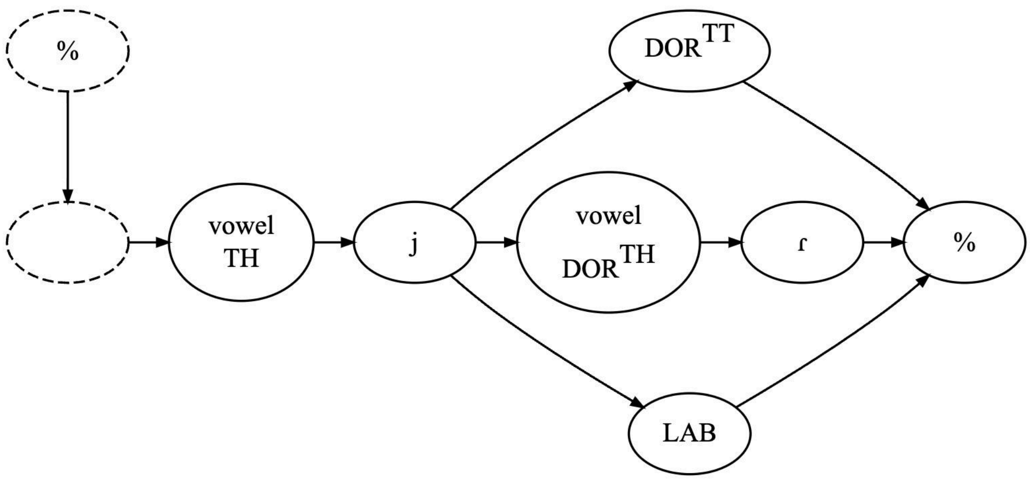
(82) *Graph representation of menkıbe [menkɯbe] 'legend'*



Because all of the vowels in *menkıbe* [menkɯbe] 'legend' match for rounding (all non round), there is one OP^{LAB} event in the rounding harmony stream which scopes over the entire word.

The 'exceptional' piece of this example is encoded in the frontness stream. First, there is a TT event which scopes over only the first vowel, /e/. Next, there is a DOR^{TT} event which scopes only over the second vowel *ı* /ɯ/. Lastly, there is another TT event which scopes only over the final /e/. This sequence of events in the autosegmental frontness tier represents the 'exceptional' front/back/front pattern displayed in the vowels of the word *menkıbe* [menkɯbe]. This process of adding an autosegmental edge to account for the 'exceptional' frontness or rounding can be used to account for all instances of invariable vowels. As a further example, consider the graph for present progressive suffix *-Iyor* /-Ijoɾ/ below. This suffix has an invariable [o] vowel in its second syllable which does not undergo harmony.

(83) *Present progressive suffix with immutable vowel -Iyor /-Ijoɾ/*



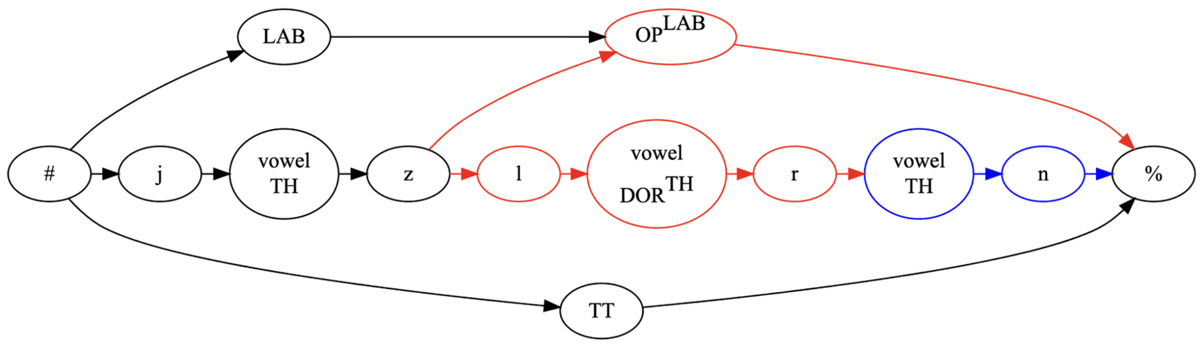
This suffix has LAB and DOR^{TT} autosegmental edges which scope over the second vowel, capturing the fact that the second vowel of this suffix is always back and rounded. If harmonizing suffixes are added after they will be under the scope of these edges, meaning they will be back, and if they are high, round.

This captures the expected pattern since TVH is opaque, as discussed earlier (*yürüyorum* [jyr-yjor-**u**m] 'I am walking', cf **yürüyorüm* *[jyr-yjor-**y**m]). Note also that the first vowel has no autosegmental edges scoping over it, indicating that it will harmonize for both frontness and rounding with the preceding vowel, as is the case for high harmonizing vowels.

3.2.2 Addressing the elephant in the room: there are no 'exceptions'

The EFP representations of the exceptional examples shown above are perhaps most remarkable in just how *un*remarkable they are. In fact, creating the phonological representation for *menkıbe* [menkɯbe] did not require use of any diacritics or new structures. Recall the example of *yüzlerin* [jyzlerin] 'your (sg) faces' from earlier, which shows rounding opacity:

(84) *yüzlerin [jyzlerin] 'your (sg) faces'*



In the above example, the first suffix introduces an OP^{LAB} into the rounding stream which blocks the spread of LAB over the second suffix. This is an exactly analogous structure to that which represents the 'exceptional' vowels in *menkıbe* [menkɯbe]; the representation is the same, but just uses frontness (TT and DOR^{TT}) rather than rounding

(LAB and OP^{LAB}). Rounding opacity is *not* an exception to TVH; it is a natural consequence of the rules of TVH. What we have done here is represent so-called exceptions to TVH in the manner as TVH itself. These 'exceptions' are not actually exceptional at all, but a natural consequence of structures already present in TVH.

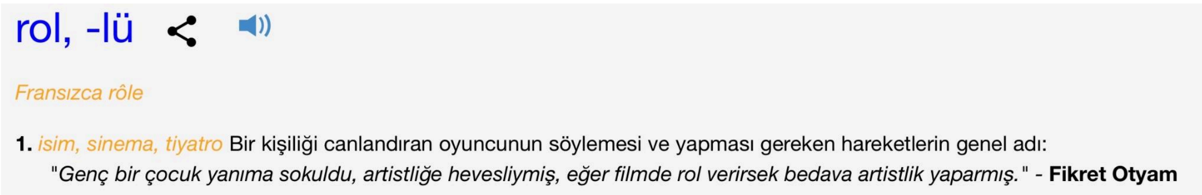
As mentioned prior, there is a lively debate about whether, and to what extent, TVH is active within roots. The answer given by this representation is that TVH is active within roots, for those roots that are marked for harmony. Roots are marked for harmony by having only a single autosegmental feature event in each of the rounding and frontness streams. Other roots are marked (via additional autosegmental events) not to conform to TVH. This type of representation is therefore free from the need to view TVH in a purely binary fashion; there is no need to determine if TVH 'always applies', 'always applies but only to certain vowels' or 'never applies' within roots. And, crucially, this representation does not make use of any kind of special diacritic mark for those roots which do not conform to TVH- it only makes use of structures which are already necessary to represent TVH, absent any so-called exceptions.

3.2.3 Exceptional front vowel suffixes

In addition to invariably disharmonic vowels, there are also a number of back vowel roots which exceptionally take front vowel suffixes. The prevailing account for these words is that they have front final consonants, which causes the suffix vowel(s) to surface as front (see Clements & Sezer 1982, Kornfilt 1999, Nevins 2010, Özçelik 2024, among others). As mentioned in the literature review, this could be a plausible explanation for those words which end in /k l/, which have front and back allophones in Turkish. However, further complicating this explanation are words which take exceptional front vowel suffixes that do not end in /k l/ (such as *saat* [sɑːt] 'clock' or *harf* [hɑɾf] 'letter').

The source of these exceptional front vowels is not the consonants themselves. It does not seem that native Turkish speakers are able to discern, from hearing a consonant on the root form alone, whether a back vowel word will take a front vowel suffix. Evidence for this can be found in dictionaries of Turkish. Take, for example, the entry for the word rol /ɾol/ 'role' in two different Turkish dictionaries, shown below in (85) and (86).

(85) *Entry from Türk Dil Kurumu Sözlükleri showing exceptional front vowel*



rol, -lü

Fransızca rôle

1. isim, sinema, tiyatro Bir kişiliği canlandıran oyuncunun söylemesi ve yapması gereken hareketlerin genel adı:
"Genç bir çocuk yanıma sokuldu, artistliğe hevesliymiş, eğer filmde rol verirsek bedava artistlik yaparmış." - **Fikret Otyam**

(86) *Entry from Türkçe Sözlük (Dil Derneği) showing exceptional front vowel*

â ç ğ ı ö ş ü

rol

ile başlayanları

sözcüğünü

ARA

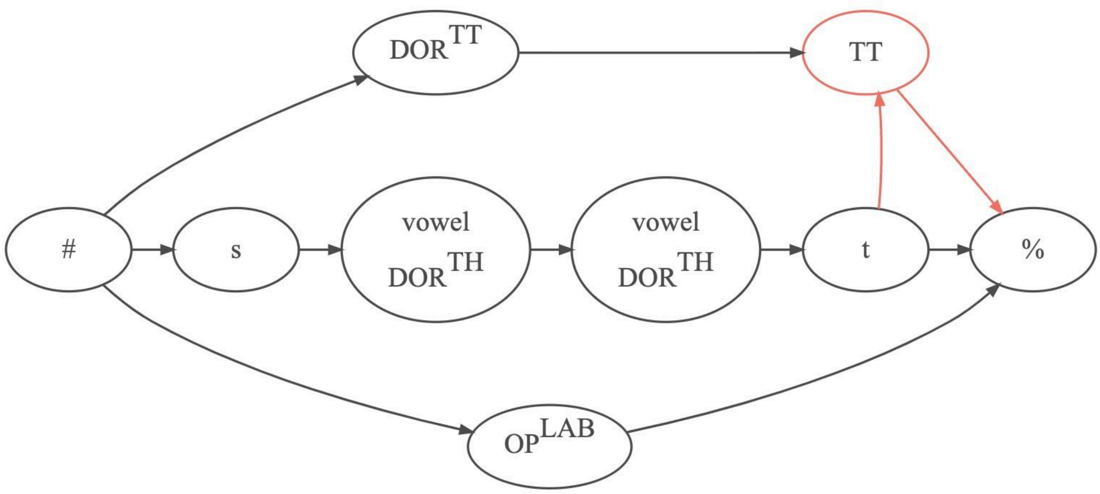
GERİ

rol, -lü *a. Fr.*
1. *tiy. ve sin.* Bir kişiliği canlandıran oyuncunun söylemesi ve yapması gereken hareketlerin genel adı.

These dictionaries are not tools made for Turkish learners such as a Turkish-English dictionary, they are monolingual Turkish dictionaries aimed at native Turkish speakers (or at least, those who speak Turkish quite well). Notice that both of these dictionaries include a small extra note after the word, *-lü*. This note does not appear on most entries in the dictionary. Its purpose is to show the reader that when suffixed with a high harmonizing vowel this word will take the front rounded allophone *ü* /y/ rather than the back rounded allophone /u/. If Turkish speakers were able to easily and reliably tell from the /l/ in the bare form alone that this word takes exceptional front vowel suffixes, no extra note would be necessary. There is no need to note, for example, that a word with a final front vowel ought to take a front vowel suffix- speakers are simply able to intuit that from the sound of the word.

Back vowel words which take front vowel suffixes can be represented in a similar manner to the invariable disharmonic vowels that were examined. As before, disharmony will be represented by adding more events into the autosegmental stream. In this case, a TT event is attached to the end of the back vowel root word. See the example below which shows an EFP graph of the word *saat* [sɑ:t] 'clock', a back vowel word which takes front vowel suffixes.

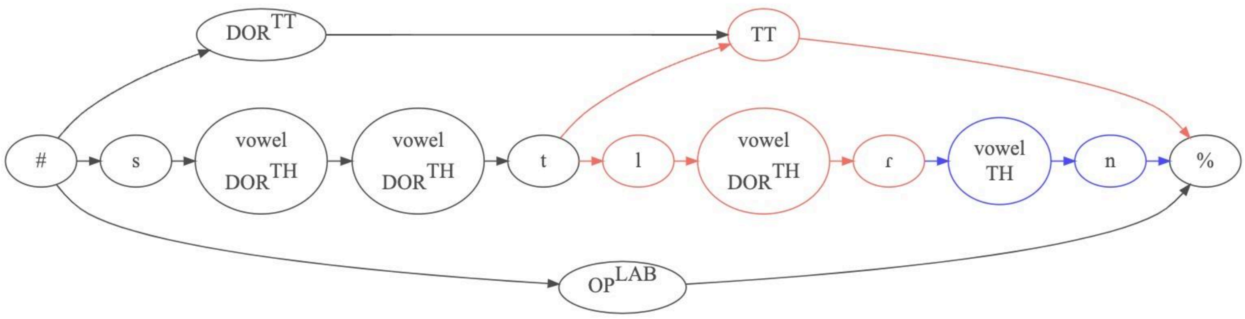
(87) *Graph of a root which takes exceptional front vowel suffixes saat [sa:t] 'clock'*



There is a TT event (red) following the word final /t/ and preceding the % (sink) node. Note that this is very similar to the second representation proposed by Papillon (2020) (see section 2.3.6 of the literature review). Unlike the previous exceptions, this additional autosegmental feature event does not scope over any vowels in the bare form. This is why there are no front vowels in the root word. However, the presence of the TT event may have some small articulatory consequence, as it may cause the /t/ to be articulated in a slightly different way, with a slightly more forward tongue position. However, since /t/ has no corresponding 'front' allophone, these differences are relatively minor (as observed in Canalis & Dikmen 2020).

When suffixes are added to *saat* [sɑ:t] shown in (87) above, they are under the scope of the TT event in the autosegmental frontness stream, yielding front vowel suffixes. The example *saatlerin* [sɑ:tlerin] 'your sg clocks' is shown in (88) below. The plural suffix *-lAr* is shown in red and the second person singular possessive suffix *-In* is shown in blue. Note that **saatların* *[sɑ:tlarɯn] with back vowel suffixes would be incorrect for this form.

(88) *saatlerin [sa:tlerin] 'your sg clocks'*



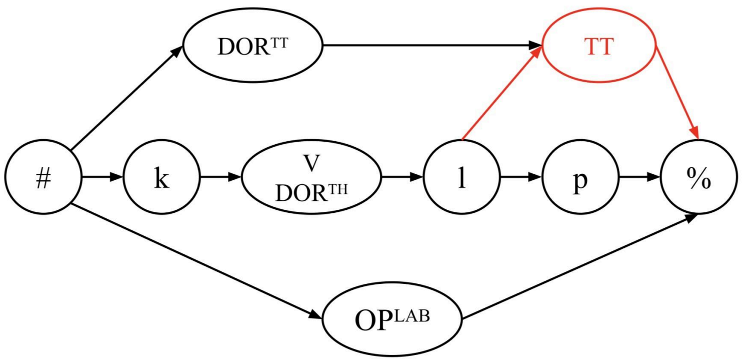
Clements & Sezer (1982) and Canalis & Dikmen (2020) (among others) have claimed that a palatal, or [-back] /t/ or /ɾ/ phoneme can cause back vowel roots to take exceptional front vowel suffixes. The important piece of this account is that there is no de-autosegmentalized consonant phoneme which marks a root for front suffixes. There may be (and we would expect this) a small phonetic consequence of the additional TT event (see Canalis & Dikmen 2020). However, this phonetic difference in the /t/ would not be the phonological *cause* of the front vowel suffixes, but rather a minor phonetic consequence of it. The actual cause, phonologically, is a TT event in the frontness harmony stream which hangs off the end of the root.

As mentioned before, it has been noted that many roots which take exceptional front vowel suffixes are loanwords with a final or near final palatal /k g l/ (See Clements & Sezer 1982 and Göksel & Kerslake 2006). In cases of root-final /k g l/, there can be a small perceptible difference for roots which take front vowel suffixes, even in the bare form, because these phonemes have front allophones. It makes sense then that the majority of these 'exceptions' would end in front palatal consonants. The 'hanging' TT events in the phonological representation are likely much easier to acquire when they are perceptible on the root word in some way. Recall from the literature review, for example, that Lewis' (1967) grammar predicted that the words *rol* [ɾol] 'role', *idrak* [idɾɑk] 'comprehension', and *sanat* [sanat] 'art' would eventually be regularized to take back vowel suffixes. Of these three, the only word that has been regularized over the last sixty years is *sanat* [sanat], a word which ends in /t/. Two of the three other words which still take front vowel suffixes are, *rol* [ɾol] & *idrak* [idɾɑk], end in /l/ and /k/ respectively- and both of these phonemes have palatal allophones.

Using an additional TT event simultaneously avoids the need to posit potential marginal palatal consonants (palatal /t/ and palatal /ɾ/) and provides an explanation for the fact that many of back vowel words which take exceptionally front vowel suffixes end in front allophones of the consonants /k l/.

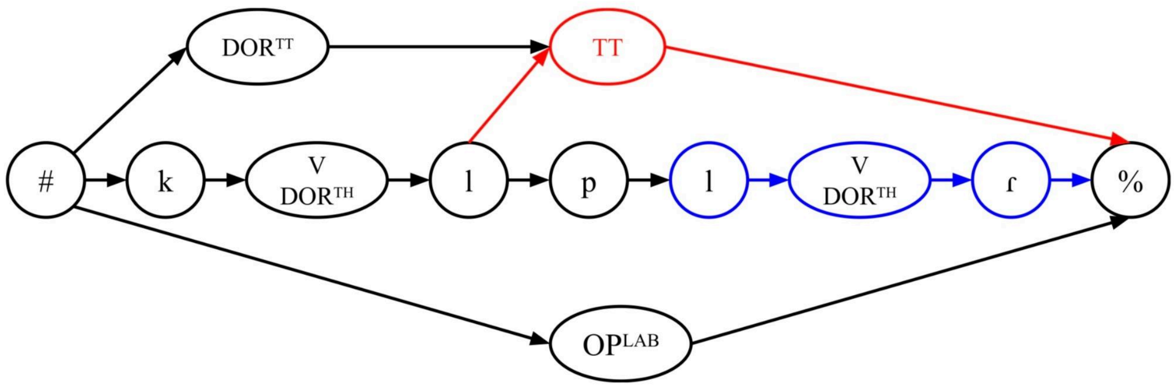
See the example of *kalp* [kɑlp] 'heart' below, which is another back vowel root which takes from vowel suffixes. The /l/ is fronted, pronounced as [l̪], a consequence of the 'hanging' TT event, which is perceptible even in bare form (see Clements & Sezer 1982).

(89) *kalp [kalp] 'heart'*



When the root is suffixed, the TT edge scopes over the new affix, causing it to surface as front, in the same manner as the earlier example. The graph below shows the root *kalp* [kɑlp] 'heart' with the low harmonizing suffix *-lAr* /-lAɾ/ added to it (blue).

(90) *Representation of kalpler [kalp-ler] 'hearts'*

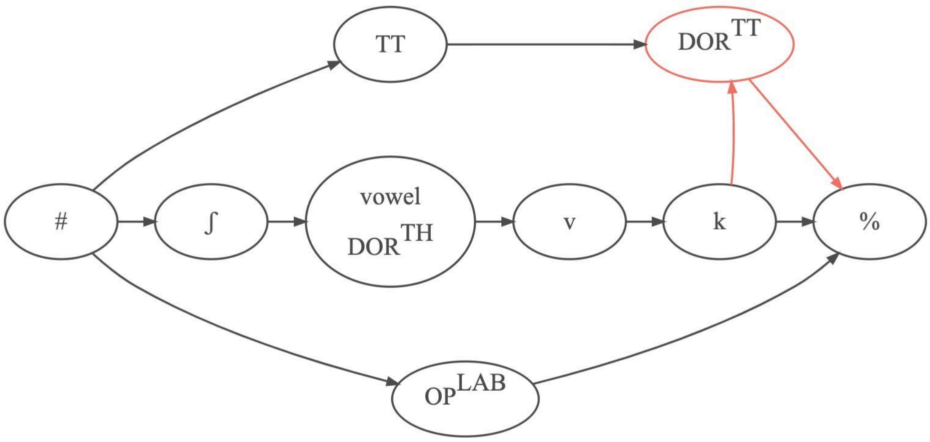


The low harmonizing suffix vowel in *-lAr* /-lAɾ/ falls under the autosegmental TT event which hangs off the right edge of the root, causing the vowel to be pronounced as [e]. Since TVH is opaque, the earlier [a] vowel in the root *kalp* [kɑlp] does not affect the quality of the suffix vowel.

3.2.4 Exceptional back vowel suffixes

As mentioned earlier, Clements & Sezer (1982) note that for a small number of older speakers, there may be front vowel words which take back vowel suffixes (this type of exception is also mentioned in Lewis 1967). While we tentatively say that these forms are likely obsolete, we will still account for them here. These can be encoded in an analogous manner to the previous front suffix vowel examples. However, rather than a ‘hanging’ TT event, these forms would have a ‘hanging’ DOR^{TT} event. See the example of *şevk* [ʃevk] ‘desire’ shown below. A DOR^{TT} event (red) follows the final consonant /k/, and does not scope over any vowels in the bare form.

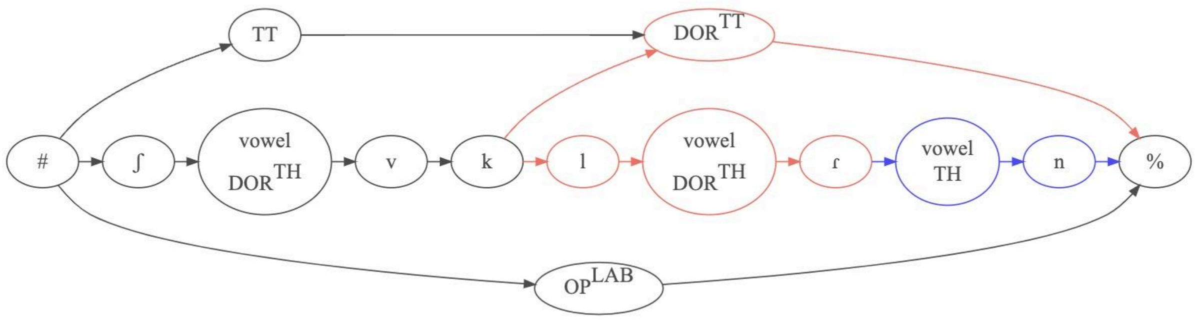
(91) *Root which takes exceptional back suffix vowels şevk [ʃevk] ‘desire’*[19]



When any suffixes are added, the DOR^{TT} edge then scopes over them, yielding back vowel suffixes, as shown below.

[19] As noted earlier, it seems that this form has been completely regularized and is now pronounced (and listed in dictionaries) as *şevki* [ʃevki], along with other ‘exceptions’ of this kind. This word is listed in Clements & Sezer 1982 as one of the few which is suffixed with a back vowel by a small number of speakers, so I provide an analysis.

(92) *Exceptional back suffix vowels şevkların [ʃevk-lar-ɯn] 'your (sing) desires'*



3.2.5 Again, there are really no exceptions

It is worth noting that we have done nothing in the preceding subsections which we have not already done. Our representation for roots which take exceptional front or back suffix vowels like that of roots with invariable vowels which did not conform to TVH, is exactly analogous to rounding opacity. To encode roots which take exceptional front or back suffix vowels, we introduce a TT or DOR^{TT} edge which does not scope over any vowels in the root. This is the same representation as that of roots with invariable vowels which did not conform to TVH, just shifted to the right so that the additional edge does not scope over any vowels in the root.

3.2.6 The case for a principled gap?

There are some combinations of vowels which are not attested within Turkish words, namely sequences involving *ı ü* /ɯ y/ and *ı ö* /ɯ œ/ (in either order). It is possible that there is some special status afforded these combinations, and they are truly prohibited in the phonology of TVH. One possibility is a disallowing of TT and LAB edges to begin or end at the same point within a word.

However, there is another significant factor, also discussed by Kabak (2011). Disharmonic *ü ö* /y œ/ often occur in loanwords from French, a language with front rounded vowels, e.g.:

(93) *ö ü /œ y/ in loan words of French origin*

	Turkish	*Transcription*	*French*	*English*
a.	menü	/meny/	menu	‘menu’
b.	pandül	/pɑndyl/	pendule	‘pendulum’
c.	modül	/modyl/	module	‘module’
d.	asansör	/ɑsɑnsœɾ/	ascenseur	‘elevator’
e.	şoför	/ʃofœɾ/	chauffeur	‘chauffeur’
f.	fritöz	/fɾitœz/	friteuse	‘deep fryer’

In addition, schwa in English is often rendered in Turkish loans as the high back unrounded vowel /ɯ/. This is the source of a number of exceptional ı /ɯ/ vowels:

(94) *Loanwords of English origin with dishomic ı /ɯ/ vowels (partially from Kabak 2011:14)*

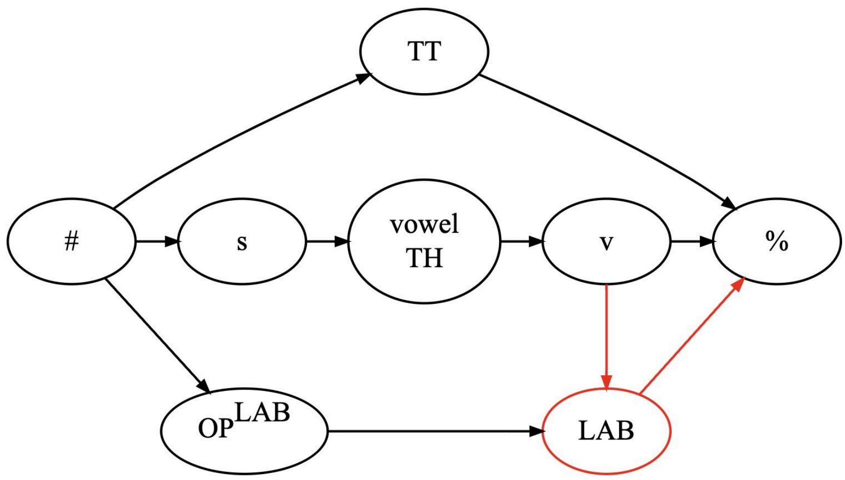
	Turkish	*Transcription*	*Translation*
a.	blendır	/blendɯɾ/	‘blender’
b.	rodstır	/ɾodstɯɾ/	‘roadster’
c.	printır	/pɾintɯɾ/	‘printer’
d.	tikıt	/tikıt/	‘ticket’ (esp in the context of ‘ticket restaurants’)

While there are some loanwords which are not from French that contain disharmonic *ü ö* /y œ/ (e.g., *zatürre* /zatyɾe/ ‘pneumonia’ from Arabic), and there are some some loanwords with disharmonic *ı* /ɯ/ which are not from English (e.g., *kırlent* /kɯɾlent/ ‘pillow’ from Italian), we can see from this information that there is a general pattern forming. There seems to be a paucity of words from donor languages which would contain *both* front rounded vowels and schwa-like vowels which would be rendered in Turkish as *ı* /ɯ/. French, while it has many front rounded vowels, does not have a schwa-like vowel. English, on the other hand, has a schwa vowel, but does not have front rounded vowels. So perhaps, as Kabak (2011) suggests, this gap has less to do with some phonological rule against such vowel combinations, and more to do with a lack of donor language input with such vowel combinations.

3.2.7 Why are there no hanging edges for rounding?

There are back vowel roots which take exceptional front vowel suffixes, and (at least at one point in time, for some speakers) front vowel roots which take exceptional back vowel suffixes. However, there are no attested unrounded vowel roots which take rounded suffix vowels, or vice versa. As an example, suppose that there is some word *siv* [siv] with an unrounded vowel, which would take rounded suffixes, as in *sivüm* [siv-ym] 'my siv'. We could represent this exception using a hanging labial event following the final consonant, much in the same way as we accounted for exceptional front or back suffixes using a hanging TT or DOR^{TT} event:

(95) *Nonce word siv which would take exceptional round suffix vowels*



It is not immediately clear why roots like [siv] above would be absent. Like the front and back allophones of /k g l/, Turkish also has labial allophones of /f v/ which occur adjacent to rounded vowels. For some Turkish speakers, /f/ → [Φ] adjacent to rounded vowels (Göksel & Kerlsake 2006). In addition, /v/ → [ω] between two vowels, where at least one vowel is round. Lastly, /v/ → [β] adjacent to a round vowel in non-intervocalic environments (Göksel & Kerlsake 2006). In addition, we do see exceptionally rounded or unrounded invariable vowels within morphemes (for example, in roots *kuzin* [kuzin] 'cousin' and *virüs* [viɾys] 'virus', and in the present progressive suffix *-Iyor* /-Ijoɾ/).

One potential reason for this gap is difficulty in acquisition. Recall that non-high suffix vowels block rounding harmony (rounding opacity) via an autosegmental OP^{LAB} stream. So, rounding harmony from the base can only spread to high suffix vowels which occur *before* any non-high suffix vowels. This would cause any hanging LAB or OP^{LAB} event imperceptible making it much more difficult to acquire. See the example below with the nonce word *siv*, which takes an exceptionally rounded suffix vowel.

(96) *Nonce word which takes disharmonic rounded suffix*

	Nonce	*Transcription*	*Gloss*
a.	siv	[siβ]	nonce word (noun)
b.	sivü	[siω-**y]**	nonce word+accusative
c.	sivleri	[siβ-ler-**i]**	nonce word+plural+accusative (exceptional rounding is blocked)

This is in contrast with the existing roots that take exceptional front vowel suffixes. Since all harmonizing suffixes undergo frontness harmony, non-high suffix vowels do not block the spread of the hanging span. For instance, compare the form in example (35c) to *saatlerin* [sɑːt-ler-in] 'your (sing) clocks'. In the latter, the TT autosegmental edge is able to scope over all the suffixes, and is not blocked by a non-high suffix vowel.

3.3 Why nonconformity is dispreferred

While there are plenty of words that do not conform to TVH, it is worth acknowledging that Turkish tends to exert some kind of pressure for words to conform to TVH, especially those with /y œ ɯ/. For example, as Clements & Sezer (1982) recognized, some words have regularized pronunciations, especially those words which contain /y œ/. Here are some examples of this from Clements & Sezer (1982):

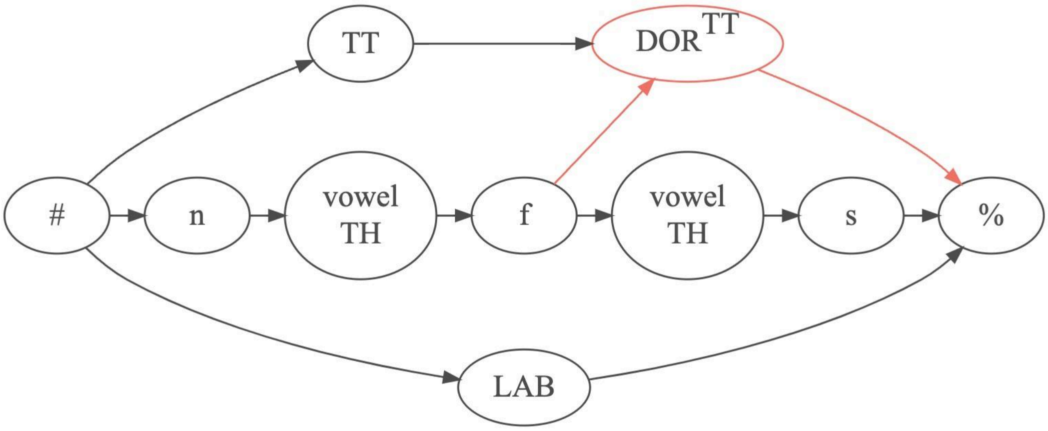
(97) *Regularization of disharmonic ö ü /œ y/*

	Turkish	*Transcription*	*Translation*
a.	ş**o**för ~ ş**ö**för	[ʃ**o**fœɾ ~ ʃ**œ**fœɾ]	'chauffeur'
b.	n**ü**fus ~ n**u**fus	[n**y**fus ~ n**u**fus]	'population'
c.	kom**ü**nist ~ kom**i**nist	[kom**y**nist ~ kom**i**nist]	'communist'

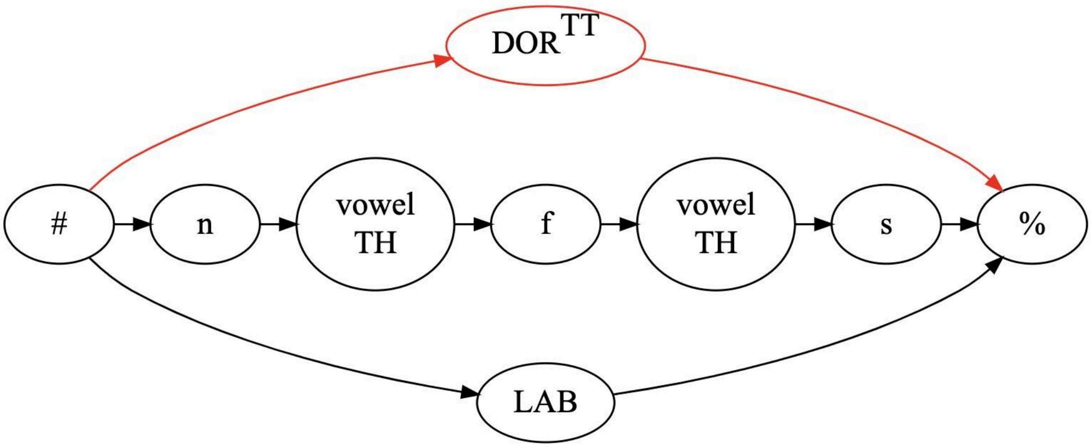
In addition, there does seem to be some pressure over time for Turkish speakers as a whole to regularize forms that do not conform to TVH, until the unregularized forms are no longer in use. Consider for example, the regularization of *sanat*, *şevk*, and *fevk*, which no longer take exceptional suffix vowels.

We can represent this pressure to regularize as a pressure to include as few autosegmental edges as possible. Speakers may 'strip off' some autosegmental edges in order to achieve this, especially in fast or casual speech. In addition, this gels well with our representations: words which conform to TVH encode less information than their nonconforming counterparts and are therefore in some sense 'easier' to represent. See the non-regularized root below *nüfus* [nyfus] 'population' in (98) which has two events in its autosegmental stream for frontness: TT followed by DOR^{TT}. Then compare the regularized form of *nüfus* as [nufus] in (99), which has only one event in the same stream: a single DOR^{TT}.

(98) *Non-regularized form of nüfus [nyfus] 'population'*



(99) *Regularized pronunciation of nüfus as [nufus]*



The structure of the regularized form in (99) has fewer events in the autosegmental frontness stream, and is in that sense a simpler representation. This simplicity may make it easier to maintain as a representation and to pronounce, and thus may be used by some speakers of certain dialects or in certain contexts (e.g., fast speech).

3.4 Loose ends: prefixes and (potential) epenthetic vowels

There still remain a few lingering issues which should be addressed concerning Turkish vowel harmony. The first is prefixes; thus far only the representation of suffixes has been laid out. Secondly, there are what have been labeled as epenthetic vowels in both onset and coda clusters in Turkish.

3.4.1 Circling back to the prefixes

The vast majority of Turkish affixes are suffixes; however, there are a handful of prefixes. Most of these prefixes are non-productive and all are loans from other languages (Göksel & Kerslake 2006). Since vowel harmony spreads from left to right in Turkish, it seems possible that prefixes would alter the vowel harmony of the root, but this is not so. These prefixes never interact with the harmony on the root. Below are some examples, taken from Göksel & Kerslake (2006:64):

(100) *Turkish prefixes*

	Turkish	*Transcription*	*Translation*
a.	**anti**demokratik	/**ɑnti**demokɾɑtik/	‘antidemocratic’
b.	**post**modern	/**post**modeɾn/	‘postmodern’
c.	**gayri**müslim	/**gɑjɾi**myslim/	‘non-Muslim’
d.	**bi**haber	/**bi**hɑbeɾ/	‘unaware’, ‘ignorant’
e.	**na**mütenahi	/**nɑ**mytenɑhi/	‘infinite’

Of these prefixes, *anti-* /ɑnti-/ is the only prefix which shows some productivity (Göksel & Kerslake 2006). There are a couple possibilities for how these prefixes are interacting with TVH. One possibility is that they are not actually parsed as affixes by the speaker, but rather as simply as a part of the root word. This is perhaps correct given the very low degree of productivity in these prefixes. In this case, we can deal with the prefixes in the same manner as a disharmonic root (see section 3.2.1).

A further possibility is that prefixes are considered by the speaker to be separate from the root, but rather than being stored as a prefix and a base, these forms are stored as compound words. Turkish, like many other languages, does not have vowel harmony across the roots in compound words:

(101) *TVH does not operate across compound words*

	Turkish	*Transcription*	*Translation*	
a.	bu-gün	[bu-gyn]	'today'	lit. this day
b.	gök-kuşağı	[gœk-kuʃɑɯ][20]	'rainbow'	lit. sky belt
c.	bilgi-sayar	[bilgi-sɑjɑɾ]	'computer'	lit. information counter

If the prefix and base is represented phonologically as a compound word by Turkish speakers, no harmony process would be expected to apply. The focus of this dissertation is not on compounds, so a representation will not be provided here for TVH in compounds. Suffice it to say that either of the two possibilities given above can adequately account for the very few Turkish prefixes.

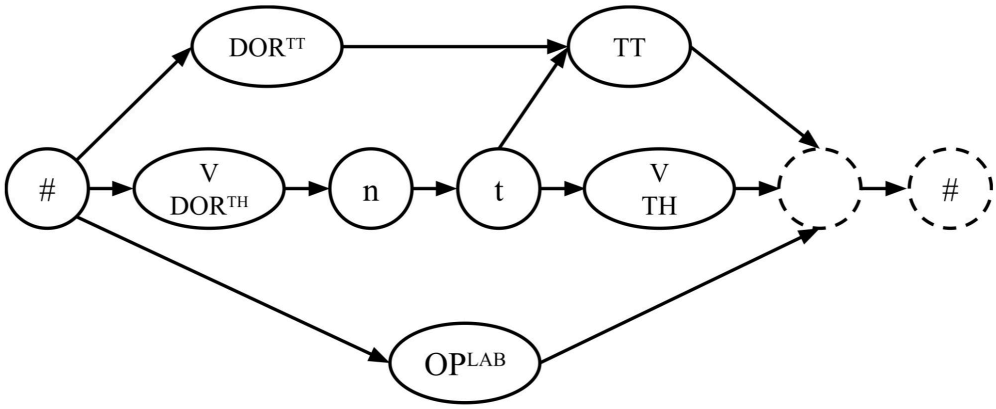
There is a third possibility, which is that speakers represent these prefixes as actual affixes with both frontness and rounding information in their autosegmental streams. Unlike suffixes, which can have an empty frontness and/or rounding stream, prefixes must have at least one event in each stream (one frontness event and one rounding event). The presence of these harmony events stops the prefixes from undergoing any kind of harmony with the root.

When a prefix is affixed to a root, the events in the harmony streams either simply slot in before the events on the root (if they are different) or combine with the existing events (if they are the same). Because the root already contains events which scope over the whole word, there is no way for the autosegmental harmony events to scope over the vowels in the word. Therefore, a prefix does not condition any harmony in the vowels of the root.

The graph in (102) below shows the EFP representation for *anti-* /ɑnti-/ from the table in (100). It comes with events in both the frontness and rounding streams.

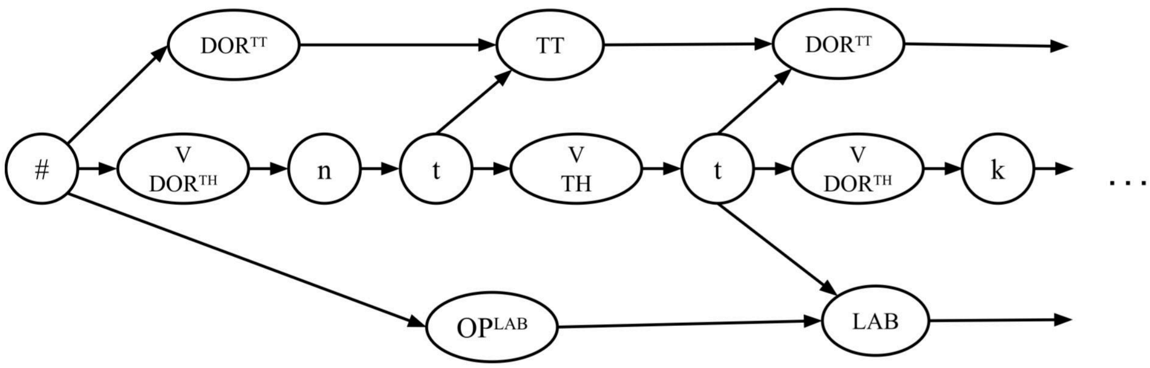
[20] Also may be pronounced as [gœk-kuʃɑː]

(102) *EFP representation of prefix anti- /anti-/*



The graph in (103) below shows the EFP representation for the beginning of the word *antitoksik* [ɑnti-toksik] ‘antitoxic’. The autosegmental harmony events from the final vowel of the prefix (TT and OP^{LAB}) do not affect the vowels of the root, because the root already has events in its harmony streams (in this case, DOR^{TT} and LAB). As a result, the harmony events from the suffix have no effect on the vowels in the root.

(103) *Prefix and beginning of antitoksik [anti-toksik] ‘antitoxic’*



This subsection has proposed several possible analyses for prefixes, which do not harmonize with the rest of the word. The first is that prefixes in Turkish are not true affixes, that they are represented as part of the root word. In this case, prefixes can be represented in the same manner as roots with disharmonic vowels. A second possibility is that words with prefixes are represented as compounds, where the prefix and root form the two parts of the compound. Vowel harmony does not spread across the roots in compounds, so no harmony would be expected to apply. Finally, there is the possibility that prefixes are indeed encoded as actual affixes. In this case, prefixes would be required to have at least one harmony event in each of their autosegmental harmony streams (frontness and rounding). This renders them resistant to undergoing harmony with the root they attach to. And, since roots in this account always have harmony information encoded with them, the harmony events of the prefix have no effect on the vowels of the root.

3.4.2 Epenthetic vowels…?

There are two environments in which epenthetic vowels have been reported to occur in Turkish. The first is the insertion of a vowel to break up consonant clusters in onsets of loanwords, which will be discussed briefly here. The second reported epenthesis is a ∅ ~ V alternation which can occur in coda clusters at the ends of words, which will be discussed in the next chapter, Chapter 4. Both types of epenthetic vowels are under debate as to whether or not they constitute true epenthetic vowels. I refer to the added vowels as ‘inserted vowels’ in this subsection for simplicity.

The first type of inserted vowel occurs to resyllabify consonant clusters in onsets. It is reported that Turkish does not allow consonant clusters in syllable onsets (Clements & Sezer 1982, Kabak 2011); however, many loanwords do contain onset clusters. In order to remedy this, a vowel is inserted either before or within the cluster. Here are a several examples:

(104) *Onset clusters in loan words*

	Turkish	*Transcription*		*Translation*
a.	tren	/tɾen/	[tiɾen]	'train'
b.	kral	/kɾɑl/	[kɯɾɑl]	'king'
c.	plan	/plɑn/	[pilɑn]	'plan'

There are mixed reports as to how often speakers pronounce these inserted vowels. Kabak (2011) reports that to pronounce a complex onset without the inserted vowel is never found in native speech and would cause a speaker to sound 'foreign'. Bellik (2018), on the other hand, states that these vowels are 'optional'. Clements & Sezer (1982), along the same lines, say that clusters are produced faithfully in careful speech, but that inserted vowels are present in more 'normal' or 'casual' speech.

In terms of whether these inserted vowels undergo vowel harmony, a few key generalizations have been proposed. Clements & Sezer (1982) claim that with a few exceptions, inserted vowels which follow velars are always back, regardless of the backness of the following vowel, eg *kredi* /kɾedi/ [kɯɾedi] 'credit'.

The reason given is that the velars at the beginning consonant clusters are uniformly back, and the epenthetic vowel undergoes harmony to match the backness of the preceding consonant. When the first consonant is not velar, Clements & Sezer (1982) state that the inserted vowel will harmonize in frontness with the following vowel, eg *spiker* /spikeɾ/ [sipikeɾ] 'speaker'.

Lastly, in terms of rounding harmony, Clements & Sezer (1982) state that inserted vowels will always match for rounding, except for those before /o/, e.g., /spoɾ/ [sɯpoɾ][21] 'sport' cf., *[supor].

[21] This word is transcribed as [sipor] by Göksel & Kerlaske (2006). Regardless, the inserted vowel [i] does not match the following vowel [o] for rounding.

Kabak (2011) corroborates this rounding harmony generalization, but is slightly more tentative, stating that inserted vowels 'may' undergo rounding harmony, and reporting some variation (e.g., *fylyrt* ~ *filyrt* 'flirt'). Kaun (1999) conducted a production experiment in which 9 native Turkish speakers pronounced 107 loan words with onset clusters. Results indicated that while speakers uniformly produced a rounded inserted vowel before high rounded vowels /y u/, there was a high degree of variation before non-high rounded vowels /œ o/. Of the nine speakers, only two exhibited the pattern predicted by Clements & Sezer (1982) in which inserted vowels always harmonize for rounding, except for those before /o/. For two speakers, inserted vowels always harmonized, and for the remaining five, inserted vowels either sometimes or never harmonized before /œ o/.

Finally, Bellik (2018) conducted an acoustic study of inserted vowels in Cr clusters as produced by six native Turkish speakers. Both nonce words and real words were included in the study. Most productions (88%) had an inserted vowel present, but this varied by speaker. Inserted vowels were found to have a shorter duration as compared to lexical vowels. In terms of vowel quality, most inserted vowels seemed similar to [ɯ], regardless of the preceding consonant or following vowel. Based on F1 and F2 measurement taken at the vowel midpoint, acoustically, inserted vowels tended to be of intermediate quality between [i u ɯ] and were gradiently impacted by the height and backness of the following vowel. Given all of this, Bellik concluded that inserted vowels do not have the timing and articulation targets associated with lexical vowels, and are intrusive, rather than epenthetic, vowels.

Based on the information laid out in this section, I argue that this kind of inserted vowel is not an epenthetic vowel. They are too unstable and too unlike other vowels in the language to be the result of a phonological epenthesis rule. These inserted vowels can be excluded based on register, and there is interspeaker variation as to how often they are inserted (see Clements & Sezer 1982, Bellik 2018). I am therefore in agreement with Bellik (2018) who found that these are best described as intrusive vowels. These intrusive vowels do not interact with phonology, and thus do not behave like typical Turkish vowels. This would explain why they are not always pronounced, and also do not seem to interact with the harmony system like a truly epenthetic vowel would.

Finally, there is one datapoint which may seem to contradict the argument above that characterizes these inserted vowels as intrusive. This evidence comes from one instance of Turkish emphatic reduplication (TER), a form of reduplication which applies to a limited number of words, which is covered in far more detail in Chapter 5 in much more detail. TER involves reduplicating the initial onset (if present) and first vowel of a word, and adding a linker consonant between the copied part and the base. The word *gri* [gɾi] 'grey' is reduplicated as *gıpgri* [gɯ-p-gɾi] 'completely grey'. The reduplicated vowel appears to be the intrusive vowel, which is odd, since intrusive vowels should not be present in the lexicon or phonology. To my knowledge, this is the only word with a complex onset which is eligible to undergo TER. I do not believe this data point is truly problematic given the account of TER adopted in this thesis; this word will be revisited in Chapter 5.

3.5 Conclusion

The goals of this chapter have been ambitious. In short, my aim was to provide a comprehensive account of TVH which encompasses forms which conform to TVH along with those which do not. To this end, I have proposed an extension of Papillon's (2020) PROP representation of TVH, which did not include an analysis of all exceptions. My account builds on Papillon (2020) by incorporating privative, substanceful features from Avery & Idardi (2001). Additional autosegmental harmony events were used to account for both regular (in the case of rounding opacity) and irregular forms (in the case of immutable vowels and roots which take exceptional front or back vowel suffixes). The main appeal of this approach is that it provides one unified representational structure for all Turkish morphemes, without needing to make use of any kind of exception diacritic.

In addition to accounting for the bulk of 'exceptions', several smaller issues were covered. First, Turkish has a small number of loan prefixes, and these do not conform to TVH. In addition, Turkish has so-called epenthetic vowels in both onset and coda clusters, some of which seem to conform to TVH and others of which do not. I have laid out EFP representations for the former in this chapter. The latter potentially epenthetic vowel is the subject of the following chapter.

4 Vowel Alternation in Coda Clusters

A limited set of bisyllabic roots in Turkish undergo a ∅ ~ high vowel alternation in their final coda clusters. Several examples are given below which were also given in section 2.4.1 of the literature review:

(38) *Examples of ∅ ~ high vowel alternation in coda clusters*

	Bare Form	*Vowel-Initial Suffix*
a.	vak**i**t [vɑ.k**i**t] 'time'	vaktim [vɑk.t-im] 'my time'
b.	boy**u**n [bo.j**u**n] 'neck'	boynum [boj.n-um] 'my neck'
c.	sab**ı**r [sɑ.b**ɯ**ɾ] 'patience'	sabrım [sab.ɾ-ɯm] 'my patience'

This alternation has been characterized as vowel epenthesis which repairs illicit coda clusters (Lewis 1967, Kabak 2011, Ketrez 2012, Özçelik 2024). However, there are some issues with this approach which were discussed in the literature review. The main complication to the epenthesis approach is that epenthesis appears to occur in some words even when coda clusters are licit, see below for several examples of such words, a subset of the examples shown in the literature review in 2.4.2.

(105) *Comparing coda clusters with inserted vowels to those without (Bacanlı 2020:27-28)*

	Bare Form		*English*	*Suffixed Form*		*English*
a.	kasıt	[kɑ.**s**ɯ**t**]	'intention, purpose'	kastım	[kas.t-ɯm]	'my purpose'
	üst	[y**st**]	'upper plane'			
b.	lahit	[lɑ.**h**i**t**]	'sarcophagus'	lahdim	[lah.d-im]	'my sarcophagus'
	ta**ht**	[tɑ**ht**]	'throne'			
c.	a**kit**	[ɑ**kit**]	'treaty'	aktim	[ak.t-im]	'my treaty'
	na**kit**	[nɑ**kit**]	'cash'	naktim	[nak.t-im]	'my cash'
	va**kit**	[vɑ.**kit**]	'time'	vaktim	[vak.t-im]	'my time'
	sı**kıt**	[sɯ**k**ɯ**t**]	'miscarriage'	sıktım	[sɯk.t-ɯm]	'my miscarriage'
	pa**kt**	[pɑ**kt**]	'pact, treaty'			

A second account is a vowel syncope approach, in which certain morphemes are marked for syncope of the final vowel when a vowel-initial suffix is added to the word (Erdal 2010, Bacanlı 2020). This chapter lays out a new approach to the data using the EFP framework. This new approach can be viewed as a blend of the vowel epenthesis and vowel syncope accounts.

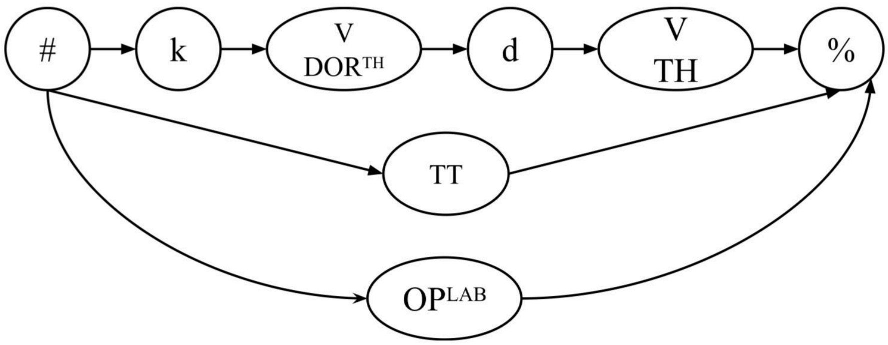
4.1 Encoding syllabification in EFP graphs

This section lays out a preliminary proposal for encoding syllabification in EFP graph structures. Because this analysis of potentially epenthetic vowels requires reference to codas and coda clusters, it is necessary to provide some information on how syllable information is represented and what constitutes a coda in this structure. Typically, in string-based phonology, syllable structure is represented using syllable trees. In this model, syllable nodes form the roots of the trees, with subcomponents of the syllable (rhyme, onset, etc.) being intermediate nodes, and phonemes being located at the leaves. An example of this kind of representation is shown below in (106):

(106) *Syllable structure tree from McMahon (2002:110, example 7)*

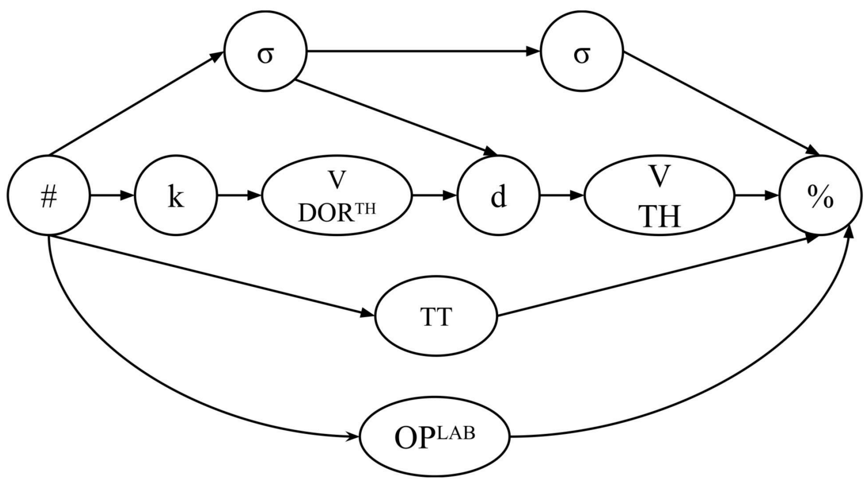
EFP uses directed graphs to represent words, which is different from the string model which the typical syllable tree is layered on top of. As such, a different system is needed to store syllables and their subcomponents in EFP. Representation of syllable structure has not yet been fully explored in EFP (see Papillon 2020 & Idsardi 2022), so I will lay out my proposed approach for Turkish here. First, see below for the EFP graph of the lexical entry for the word *kedi* [kedi] 'cat' with its vowel harmony streams for frontness and rounding.

(107) *EFP representation of lexical entry for kedi [kedi] 'cat'*



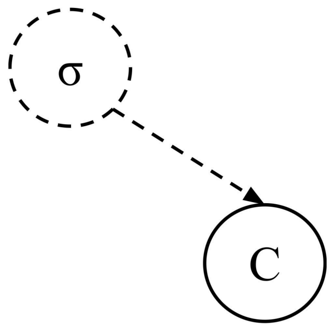
There is no syllable information stored on the lexical entry, as this word follows the regular case for syllabification in Turkish. The EFP graph below shows the representation after syllabification has occurred. Following Idsardi (2022), syllable information is stored in parallel to de-autosegmentalized information in a separate stream (see subsection 2.1.2 of the literature review for further details).

(108) *EFP representation of kedi [kedi] 'cat' after syllabification*



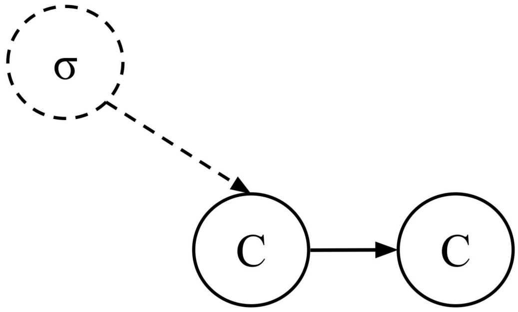
The syllable information is an autosegmental stream, and is anchored separately to the source and the sink, like the vowel harmony streams. This word has two syllables, so there are two syllable events. The word is syllabified as [ke.di], with the vowel [e] being the last member of the first syllable. Thus, the first syllable must precede the onset of the second syllable [d]. In this system, a word-medial onset can be defined as 'a consonant event which is preceded by a syllable event', as shown using an EFP representation in (109) below.

(109) *General structure of a word-medial onset*



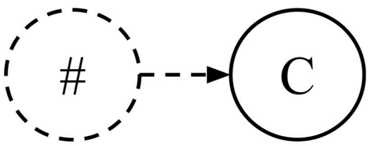
Turkish does appear to allow onset clusters, at least in some cases with word-initial onsets (see the discussion in 3.4.2 for further details). While not very common, there are some loanwords with word-medial CCC clusters which could condition word-internal complex onsets (such as *ul**tr**a* 'ultra' and *sür**pr**iz* 'surprise'). However, these may not actually be pronounced how they are written in Turkish orthography, and thus may not actually be produced with such clusters (see Meral 2023). If Turkish does allow for word-medial onset clusters the following must also be a possible onset structure for Turkish:

(110) *General structure of a word-medial complex onset*



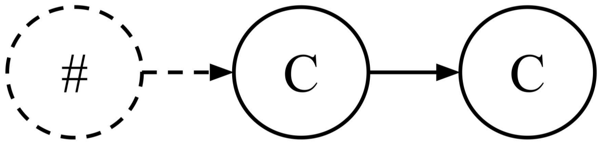
Note that the structures in (109) and (110) do not encompass the onset of the first syllable. However, this does not pose a significant issue, as the onset of the first syllable, if present, will always be the consonant in word-initial position. The structure of a word-initial onset can be characterized as ‘a consonant event which follows the source’. Here is the second structure of an onset, which captures this case:

(111) *Structure of word-initial consonant*



Finally, as mentioned above, there are some cases of word-initial complex onsets in loanwords (for example, *tren* /tɾen/ ‘train’). These may be produced with an intrusive vowel between the two consonants of the cluster (see the discussion in 3.4.2 for more information). Because intrusive vowels are not present in the phonology, Turkish must allow for complex onsets in word-initial position. The structure of a word-initial complex onset can be characterized as ‘a two sequential consonant events which follow the source’ So, the following structure must also be a valid onset in Turkish:

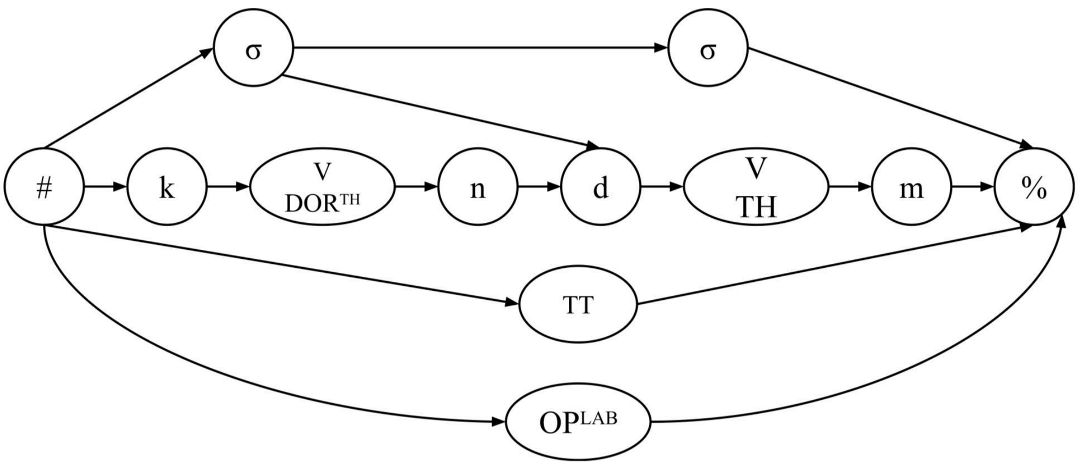
(112) *Structure of word-initial complex onset*



To clarify, Turkish phonotactics do not allow for any two consonants to be in an onset cluster together (e.g., [gf] is not attested as a coda cluster in Turkish). The generic consonant events in the general structures above are not meant to denote that any two consonants could be in an onset together; rather the purpose is to demonstrate that an onset cluster can be clearly defined in this representation. There is a separate part of the phonology which is concerned with the phonotactics of the syllables being valid, which I will not be covering as part of this work. However, syllable trees (as the one shown in 106) also do not compute phonotactics in and of themselves; rather, they show the representation after phontactic constraints on clusters have been applied. My goal here is to show how EFP graphs can represent the same structural information as a typical syllable tree.

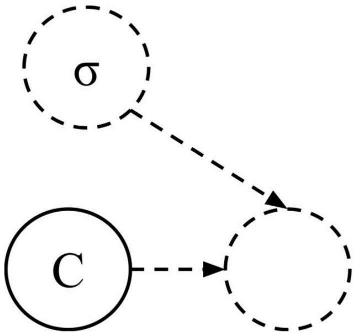
There are no codas in the example *kedi* [ke.di] 'cat' above, so we turn to a new example which does have codas: the word *ken.dim* [ken.dim] 'myself'. Below shows the EFP graph representation of this word after syllabification has occurred.

(113) *ken.dim [ken.dim] 'myself' after syllabification*



In the above structure there are two codas: [n] and [m]. Note that [n] precedes [d] and [m] precedes the sink (#). Both [d] and the sink (#) are also preceded by syllable events. A coda can be generally defined as ‘a consonant event which precedes an event also preceded by a syllable event’. This can also be characterized as: a coda is ‘a consonant event which precedes an onset or the sink’. This definition is captured in the EFP graph structure shown below in (114):

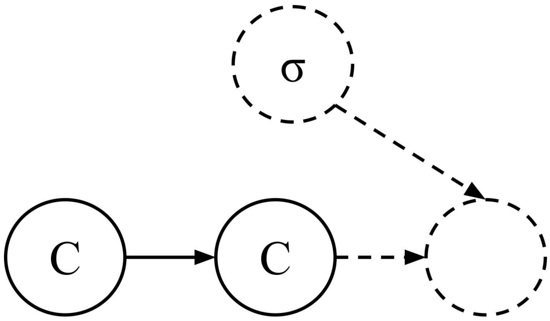
(114) *Coda structure in Turkish*



Note that there is no need to make a separate structure to capture the word-final coda, as we did for a word-initial onset. Because of the way the syllable stream is attached to words, the sink will always be preceded by a syllable event.

There is another case which the structure in (114) does not capture, however. Turkish permits complex codas consisting of two consonants. So, a coda may consist of two consonants which precede an onset. The EFP graph representation for a complex coda is shown below:

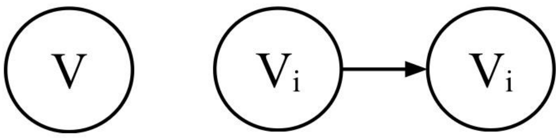
(115) *Complex coda structure in Turkish*



As before with the representation of onset clusters, this generic structure should not be taken as a claim that any two consonants can co-occur in a coda (e.g., [nl] is not attested as a coda cluster in Turkish). The purpose here is merely to demonstrate that using this representation, coda clusters can be located in the graph, once a separate part of the phonology has applied phonotactics to create syllable boundaries.

The final part of the syllable not yet explicitly defined is the nucleus. Turkish has no syllabic consonants (see Göksel & Kerslake 2006:xxiv). Therefore, in Turkish a nucleus consists of a vowel. Turkish does have lexically specified long vowels (e.g., doğru [doːɾu] 'true, towards') and in EFP these can be represented as two consecutive vowel events (see Chapter 5 for more details). Thus, a nucleus has two possible structures in Turkish: a vowel event and a two consecutive identical vowel events:

(116) *Possible nucleus structures in Turkish*



To conclude, syllable structure can be represented in EFP graphs as follows. Onsets can be characterized as up to two consecutive consonant events which follow a syllable event, or up to two consecutive word-initial consonant events (though onsets in Turkish are typically only one consonant event). Codas are up to two consonant events which precede an onset, or the end of a word. This can be captured generally as ‘up to two consonant events which precede an event also preceded by a syllable event’. Finally, nuclei are either single vowel events or two identical consecutive vowel events.

4.2 Representing vowel alternation in coda clusters

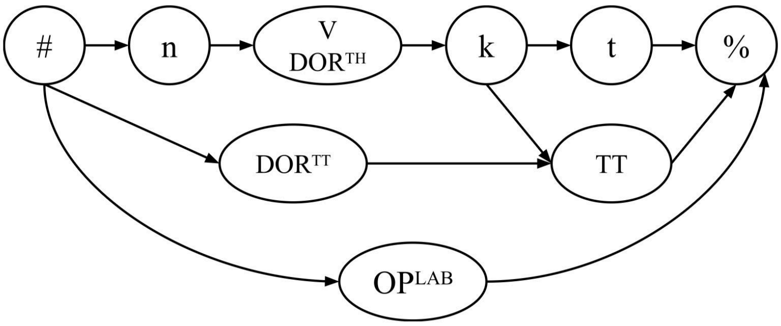
The approach given here will be a combination of the epenthesis repair for illicit clusters, and a morpholexically-conditioned syncope approach. The key piece of the representation will be a ‘dangling event’ which marks certain two-syllable words for a ∅ ~ high vowel alternation in the second syllable.

A dangling event is an event which only has one connecting edge (in graph theory vocabulary, this would be called a vertex or node with a degree of one). Crucially, the source and sink are excluded from being dangling events. The single edge can be an outgoing edge, where the dangling event precedes another event, and follows nothing. This is the kind of dangling event which will be used in this chapter. The single edge can also be an incoming edge, where the dangling event follows another event but precedes nothing. This kind of dangling event will be used to represent TER in Chapter 5. Crucially, a dangling event will not be included in the surface form while it remains dangling. In order for it to be pronounced, some rule must apply to add a precedence relationship such that it is properly sequenced with other events in its tier. For a dangling event which follows nothing, an incoming edge must be added so that the dangling event follows another event. For a dangling event which precedes nothing, an outgoing edge must be added so that the dangling event precedes another event.

4.2.1 Dangling vowel in the bare form

An EFP representation for the word *nakit* [nɑkit] ‘cash’ with its harmony spans is shown below in (117). This is a member of the limited set of words which undergoes ∅ ~ high vowel alternation in the second syllable. More will be added to this representation; what is shown below is just a starting point.

(117) *Root /nakt/ ‘cash’ with vowel harmony events*



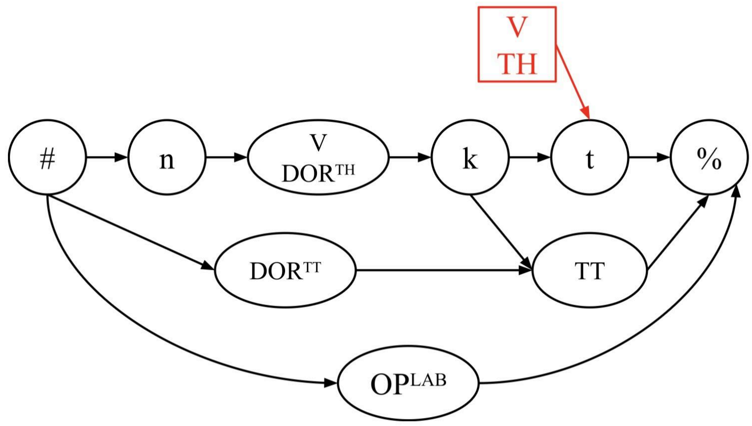
The root must consist of at least four phonemes: the consonants [n k t] and a non-high vowel. The whole word is unround, so an OP^{LAB} event spans over the entire word. The first vowel is back, hence the DOR^{TT} event which scopes over the first syllable. However, a TT event follows the [k] since both the second vowel in *nakit* [nɑkit] and any added harmonizing suffix vowels (as in nakt**i**m [nakt-**i**m] ‘my cash’[22]) are front.

Vowels inserted into coda clusters are not truly epenthetic vowels, since, as previously pointed out, vowels are sometimes inserted into licit codas. What I argue, then, is that this process is morpholexical. Certain two-syllable words are marked for ∅ ~ high vowel alternation with a ‘dangling event’ containing a vowel. This dangling event marks that the final consonant of the word cannot be in a coda cluster with the consonant before it. That is, when the final consonant of the word is in the coda, the vowel is inserted to break up the cluster. When the final consonant of the word is in an onset, the vowel will not be

[22] Some sources have this form as nakdim [nɑkdim]. Alternating consonants of this kind will be discussed later in Chapter 6.

pronounced. See below for the dangling vowel on the underlying representation of the word /nakt/:

(118) *Dangling high vowel event in the UR*

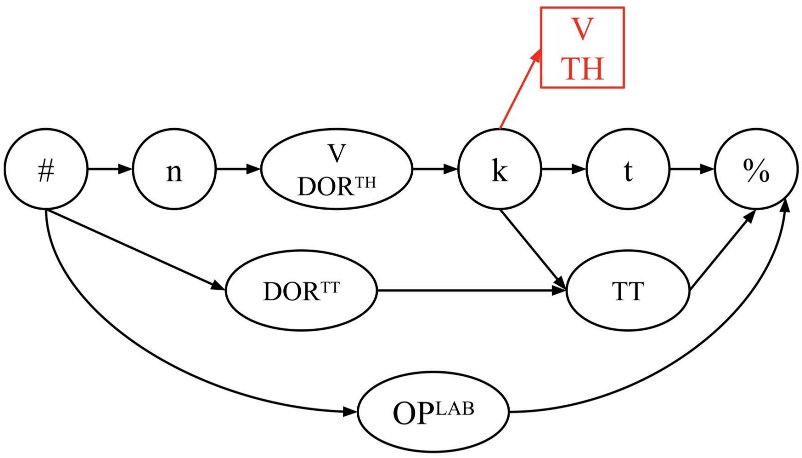


Shown above is the complete UR for this word in the lexicon. The dangling vowel event (red, square) marks /nakt/ for ∅ ~ high vowel alternation between [k] and [t]. This dangling vowel marks that the [k] and [t] in this word will never occur in a coda cluster together, even though [kt] is a valid cluster in Turkish (see *pakt* [pɑkt] 'pact', for example).

Note that the square shape and red color of the dangling vowel event is not meant to denote that a dangling event is encoded in some kind of special 'red square' event. The shape and color here are visual cues for the reader to more easily understand the structure. In this case, a dangling event is defined as an event that follows nothing and is not the source (#). There is another kind of dangling event (one which precedes nothing) which will be used in Chapter 5 for representing Turkish emphatic reduplication.

The dangling vowel event precedes the [t] to ensure that it is inserted in the correct position. That is, when [t] is in a coda, the vowel should be inserted before the [t], and not after. Since the inserted vowel should follow [k], another seemingly logical option could be to have the dangling vowel follow the [k] and precede nothing, as shown below.

(119) *Problematic dangling high vowel event following [k]*

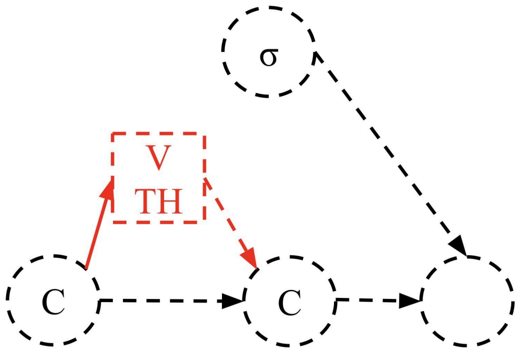


There is an issue with the above representation, however. If the vowel is only marked as following [k], it should be equally valid for the dangling vowel to surface after the [t], as in **nakti* *[nɑkti] for the bare form. In that case, while the high vowel does not immediately follow the [k], it still is pronounced after [k], satisfying the precedence relationship. In addition, sequencing the dangling vowel following the [t] would also ensure that the [k] and the [t] are not in a coda cluster together. However, this kind of insertion is never present for words with this alternation. Thus, the dangling vowel should precede [t], as in (118).

Phoneme events which are left dangling are not pronounced since they are not fully sequenced with respect to other phonemes in the de-autosegmentalized tier. A rule is necessary for sequencing the dangling vowel event with the other phonemes in the word so that it is pronounceable in the surface form when appropriate. As an example, when the [t] in /nakt/ 'cash' is in an onset, as in [nak.tim] 'my cash', then the dangling vowel should not be pronounced. Whenever the [t] would be in a coda cluster with the [k], the dangling vowel should be pronounced, as in the bare form **nakt* *[nɑkt] → *nakit* [nɑ.kit] and when suffixed with a consonant-initial suffix **naktten* *[nɑkt.ten] → *nakitten* [nɑ.kit.ten] 'from cash'. So, in general, if the consonant preceded by the dangling event is in a coda cluster, an edge should be added so the dangling vowel will be sequenced with the other phonemes and pronounced.

This can be generically stated in prose as ‘if the dangling vowel precedes a coda, sequence the vowel after the other consonant in the coda’. This structure is given perhaps in a more intuitive way by the EFP graph structure below:

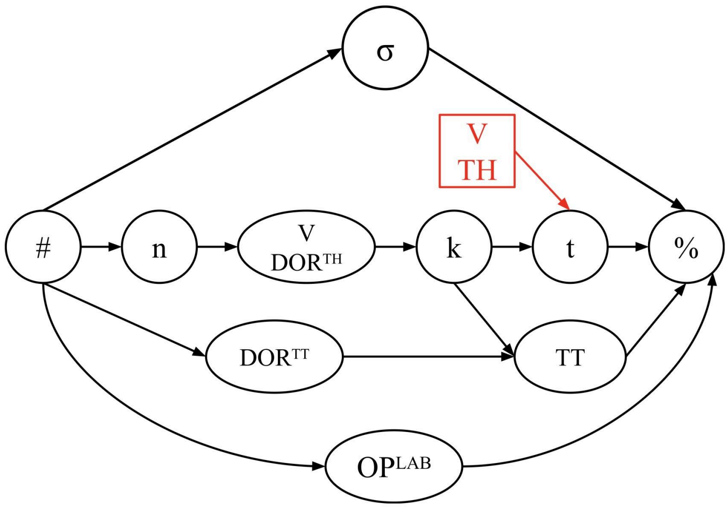
(120) *General structure for sequencing a dangling vowel in a coda consonant cluster*



The dashed parts of the graph represent parts of the structure that should be matched (e.g., the dangling vowel event must be present for the rule to apply). The solid pieces, in this case just the single edge preceding the vowel event, are what is added when the dashed structure is matched.

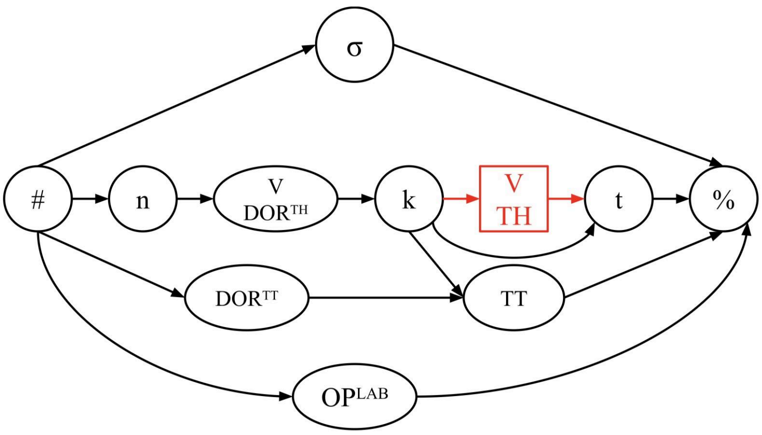
When the underlying representation for /nakt/ in (118) undergoes syllabification, the structure (121) is attained. The word is initially syllabified as including a single syllable, since [kt] is a valid coda cluster in Turkish.

(121) */nakt/ with syllabification*



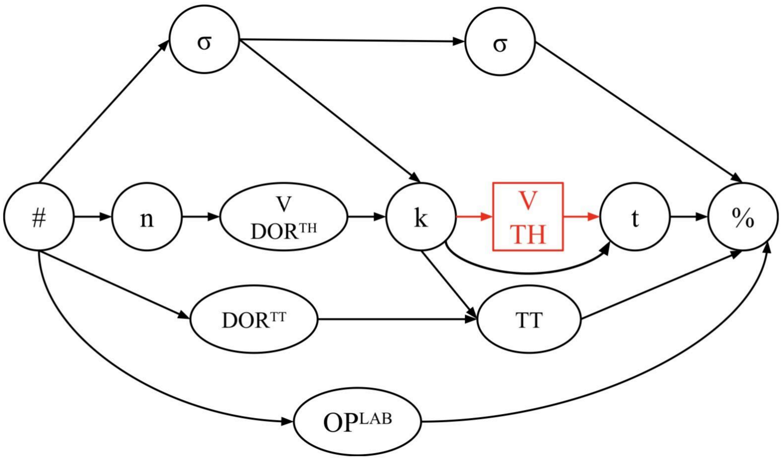
Since the [t] is in a coda cluster, the rule in (120) applies, and an edge is added such that the dangling high vowel follows the [k]. This yields the structure given in (122) below:

(122) *Edge added from [k] to [t]*



Since the high vowel is now properly sequenced with respect to the other phonemes of the word (i.e., no longer dangling), the vowel will be pronounced in the surface form following [k] and preceding [t]. This causes an issue with the syllabification, as now there are two nuclei (the two vowel events) within the same syllable. Turkish does allow for two identical sequential vowel events in the same nucleus (in the case of lexically specified long vowels), but the structure in (122) does not conform to that description as the vowels are neither identical nor adjacent to each other. Thus, the syllable structure of the word must be repaired to have two syllables, as in (123) below:

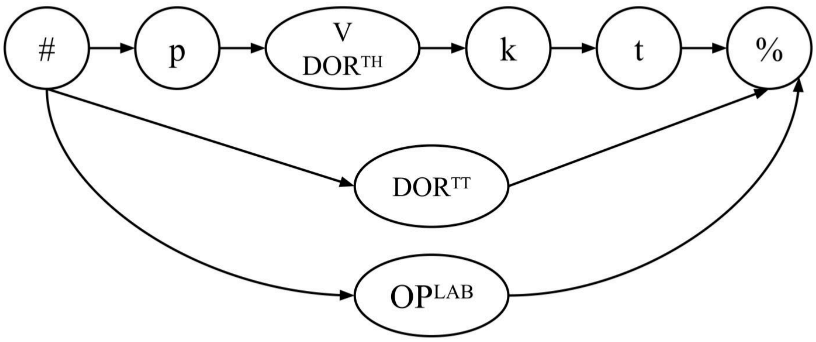
(123) *Repair of the bare form /nakt/ to form [nakit]*



The representation in (123) yields the correct surface form for the bare root *nakit* [nɑkit] 'cash'. There are two syllables, and the vowel has been sequenced in the correct place in the word.

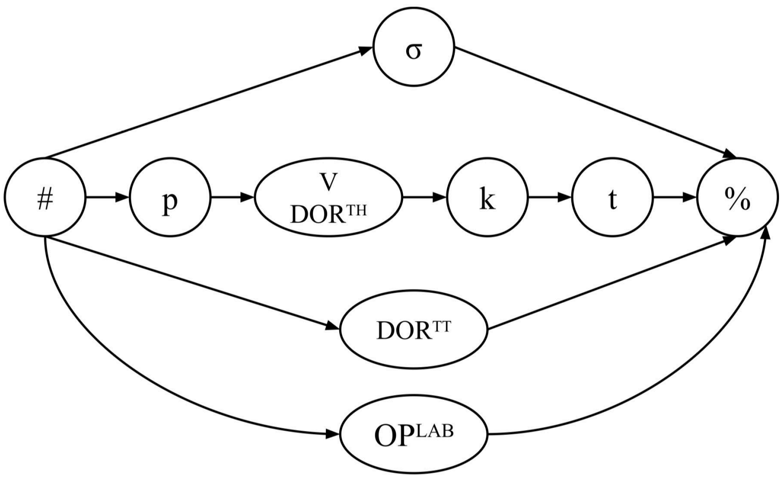
Words with identical coda clusters without high vowel ~ ∅ do not pose an issue for this analysis (as they do for the traditional epenthesis analysis). Compare the representation in (121) to the one in (124) below, which shows a graph of the lexical entry for the word word *pakt* /pɑkt/ 'pact'. Unlike *nakit* [nɑkit], this word does permit a [kt] coda cluster in the surface form. There is no dangling vowel event preceding the /t/ in /pɑkt/, so there will be no vowel inserted between the [k] and the [t]. Note that this form takes back vowel suffixes, so there is no need for a TT event following the /k/.

(124) *Representation of pakt /pakt/*



When the word above undergoes syllabification (125), only one syllable is needed, since the structural description for the vowel sequencing rule is not met (because there is no dangling vowel event). Thus, the correct surface form is attained for the representation below, the one-syllable word [pɑkt].

(125) *Syllabification of pakt /pakt/*

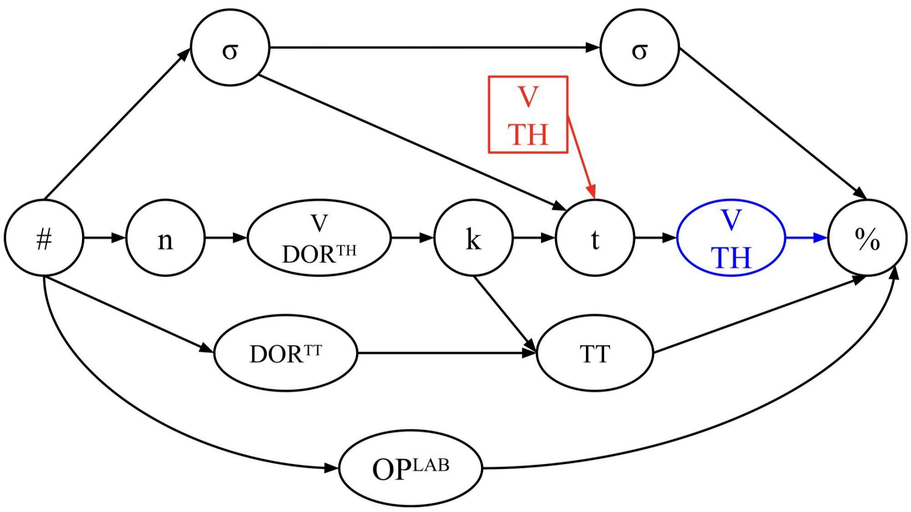


4.2.2 The dangling vowel and suffixation

The structures laid out to account for this ∅ ~ high vowel alternation also work well when suffixation is applied to words. This subsection demonstrates how the dangling vowel structure yields the correct surface form for the addition of both vowel-initial and consonant-initial suffixes. Finally, a small amount of potentially problematic data around suffixation will be discussed.

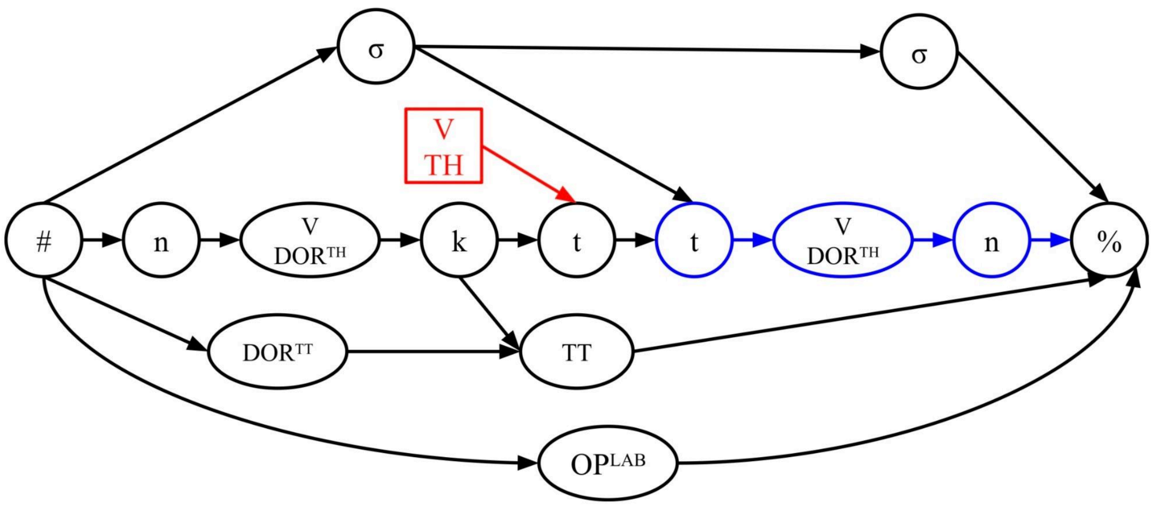
When a vowel-initial suffix is added, the situation with the dangling vowel will be different. The EFP graph in (126) below shows the root /nakt/ with the accusative suffix, consisting of a high harmonizing vowel, (blue) affixed onto it. Because of the suffix, the [t] is syllabified into an onset (represented in the graph as following a syllable node). As such, the vowel event remains dangling, since the structural description for the vowel sequencing rule is not met. Since the dangling vowel event is not sequenced properly with the other phonemes, it is not pronounced, yielding the correct surface form without the inserted vowel, *nakti* [nɑkti] 'cash-ACC'.

(126) *EFP Representation of nakti [nakti] 'cash-ACC'*



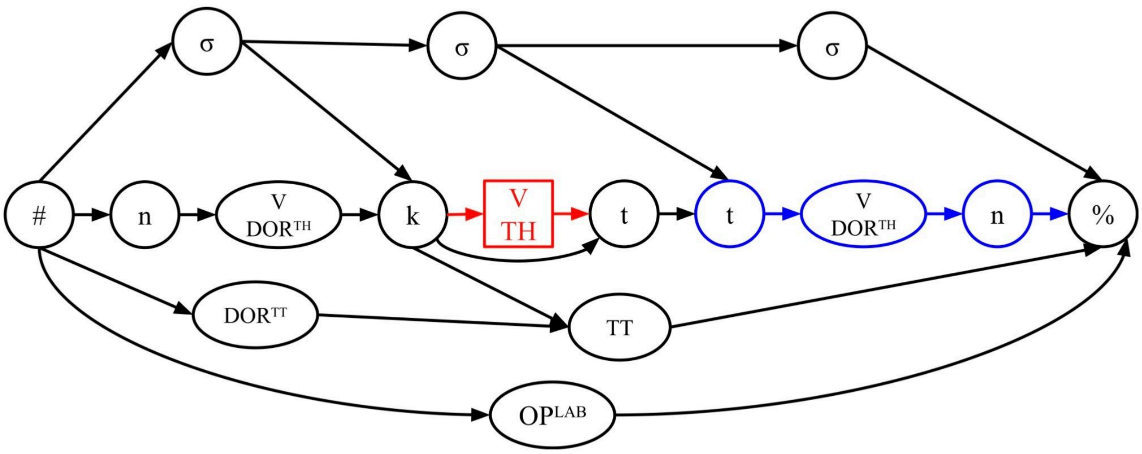
If a consonant-initial suffix is as added, then the inserted vowel should surface. For example, when the dative suffix -DAn 'from' is added, the correct surface form is *nakitten* [nɑ.kit.ten] 'from cash', cf **naktten* *[nɑkt.ten]. The representation in (127) shows the dative suffix (blue) attached to the root /nakt/ following syllabification. The dangling vowel precedes a coda [t], so the dangling vowel sequencing rule applies.

(127) */nakt-ten/ following initial syllabification*



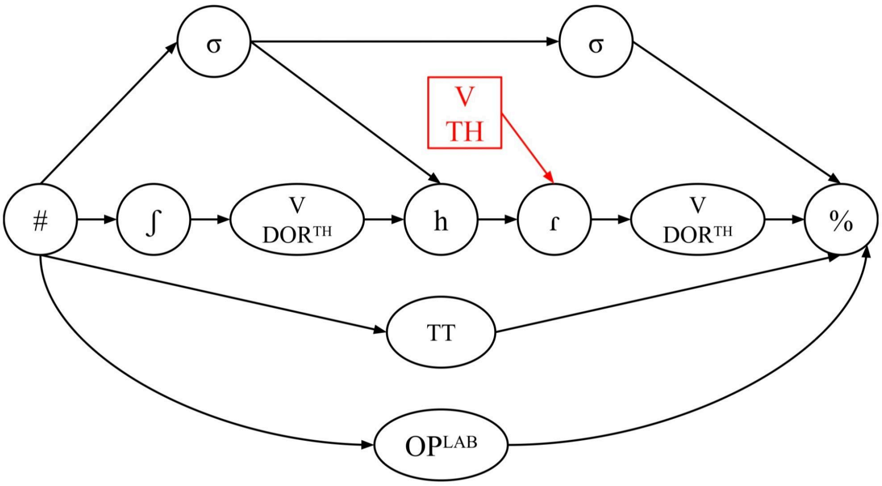
The structure in (128) below shows the form following the dangling vowel sequencing rule. Once the vowel is sequenced and pronounceable, re-syllabification occurs to accommodate the new syllable. So, the correct three-syllable surface form is attained: *nakitten* [nɑ.kit.ten] 'from cash'.

(128) *[nakt-ten] following rule application and re-syllabification*



Lastly, there is some potentially complicating data to return to regarding suffixation. Recall that it has been reported in some sources that when suffixed with a low-vowel initial suffix, the vowel in the second syllable of the root may or may not surface (Lewis 1967, Göksel & Kerslake 2006). The example given in both sources was *şehir* [ʃehiɾ] 'town'. When suffixed with a high vowel, the second vowel of the base is not pronounced: *şehri* [ʃehɾ-i] 'town-ACC' (cf., **şehiri* *[ʃehiɾi]). However, when suffixed with a non-high vowel, some speakers will include the second vowel: *şehire* [ʃehir-e] 'to town'. This is problematic for an epenthesis approach, since in both cases, the coda cluster has been resolved, and there should be no reason for the epenthetic vowel to be added in the context of a non-high vowel suffix. It is somewhat, but far less, problematic for the approach laid out here. The EFP graph in (129) below shows the representation of *şehre* [ʃehɾ-e] 'to town'.

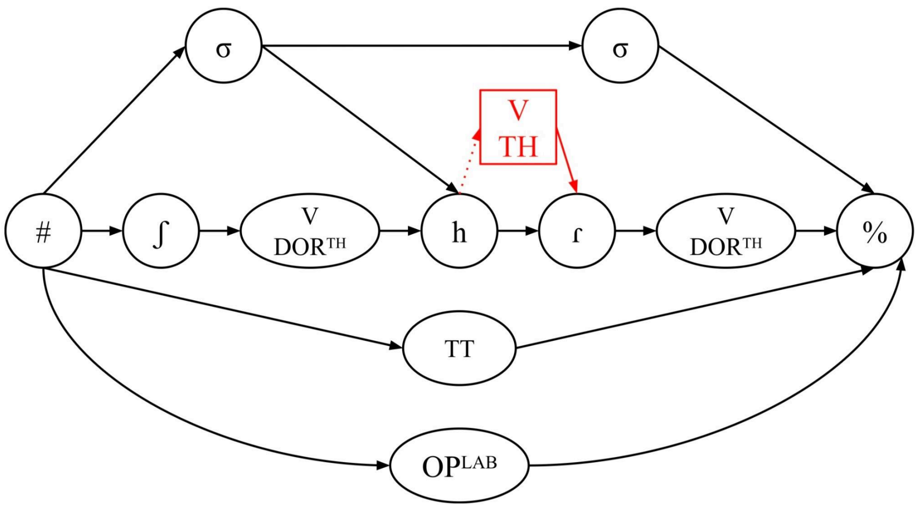
(129) *şehre [ʃehɾ-e] following syllabification*



The graph in (129) corresponds to the expected case, where the second vowel of the root is not pronounced when the *r* [ɾ] is resyllabified into an onset with the addition of a low suffix vowel. The rule for sequencing the dangling vowel event only applies when the consonant following the dangling vowel is in a coda. So, the rule is not predicted to apply in this case, and the vowel should not be pronounced, yielding *şehre* [ʃehɾ-e].

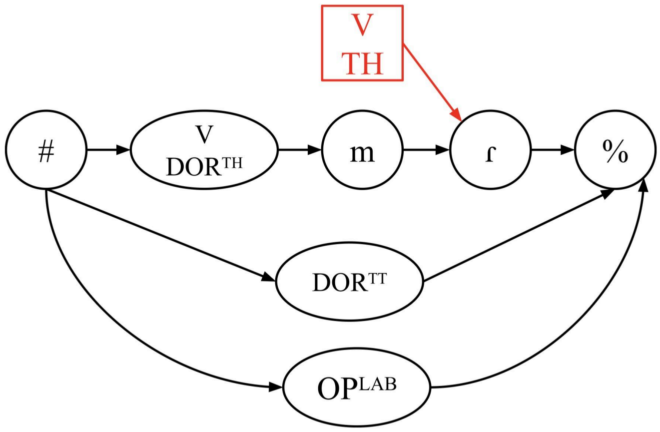
However, we know that some speakers pronounce this word as *şehire* [ʃehir-e]. So, some speakers are choosing to sequence the vowel in this case, despite *r* [ɾ] being in an onset. This corresponds to adding an edge where the dotted line is in the graph below. Since the vowel is already present in the underlying representation, this can be characterized as interspeaker variation, where some speakers sequence the vowel in a non-high vowel environment.

(130) *Dangling vowel sequenced in a non-high vowel environment şehire [ʃehir-e]*

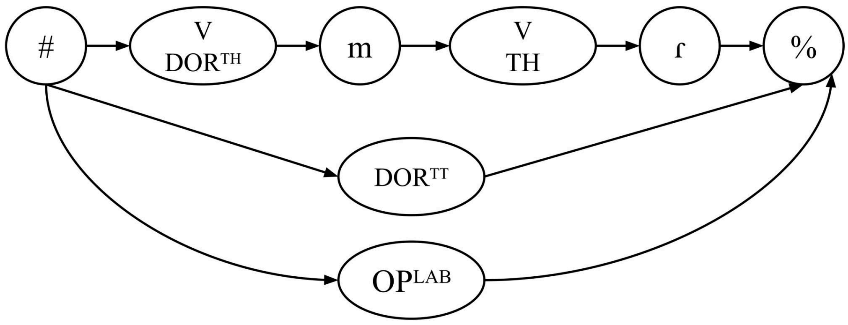


Finally, it seems that in some cases, the dangling vowel can actually become permanently sequenced with other phonemes, and cease to be dangling in the underlying representation (thus leading to no alternation). This is the case for proper nouns which are equivalent to these bisyllabic words which undergo alternation in their final coda clusters. For example, the word *ömür* [œmyr] 'lifetime' is also given as a name, *Ömür* [œmyr]. The common noun undergoes the alternation: *ömrü* [œmr-y] 'lifetime-ACC', cf **ömürü* *[œmyry]. However, the proper noun does not: *Ömür'ü* [œmyr-y] 'Ömür-ACC', cf **Ömr'ü* *[œmyry]. This points to two minimally different lexical entries: one which has a dangling high vowel event, and thus displays alternation (the common noun, see 131) and one in which the high second vowel is fully sequenced with other phonemes, and thus is always present (the proper noun, see 132).

(131) *Lexical Entry for ömür [œmyr] 'lifetime' (common noun)*



(132) *Lexical Entry for Ömür [œmyr] (proper noun)*



4.3 Conclusion

This chapter has outlined a new proposal for the analysis of what has been called vowel epenthesis (or vowel syncope) in coda clusters in Turkish using EFP graph structures. In this proposal, certain roots have alternation due to a 'dangling' vowel event preceding the final consonant. When the final consonant is re-syllabified into an onset, the dangling vowel is sequenced with respect to the other phonemes of the word and pronounced. When the final consonant stays in coda position, the dangling vowel is not sequenced and thus not pronounced. This analysis can be seen as a synthesis of the two past approaches: epenthesis and syncope. Since the vowel is in some sense ‘inserted in’ to the word when it is sequenced, this could be viewed as fundamentally similar to epenthesis. However, in another sense, the dangling vowel is present in the underlying representation but ‘left out’ of the surface form when not sequenced, similar to syncope.

5 Emphatic Reduplication

This chapter lays out how EFP can be used to represent reduplication in general, and then goes on to apply it to Turkish emphatic reduplication (TER) specifically. I argue that the choice of linker consonant for TER is fully lexicalized, and phonological representations reflect that. Finally, I lay out alternatives which account for the data on choice of linker consonant, essentially other tools we may take up instead of using phonological processes like dissimilation or identity avoidance.

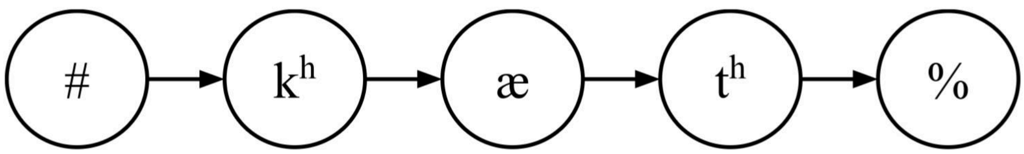
5.1 Making a loop: EFP graphs & reduplication background

This section briefly lays out how EFP can be used to represent total and partial reduplication, using examples from English and Turkish.

5.1.1 Representing total reduplication

Recall this very simple graph of the English word [k^hæt^h] cat from the literature review chapter. When it is pronounced, flow begins at the source (#), continues to [k^h], then to [æ], and so on, until it hits the sink (%), at which point traversal of the graph has finished, and the whole word has been pronounced.

(133) *Graph of [k^hæt^h] and its precedence relationships*



precedes k^h

k^h precedes æ

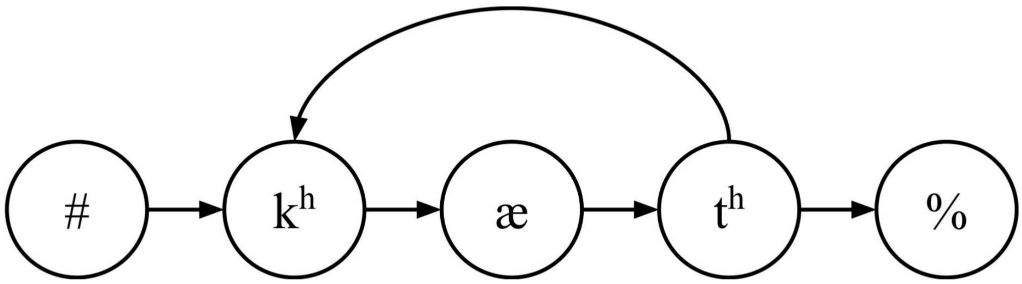
æ precedes t^h

t^h precedes %

output: [k^hæt^h]

Building on example (133) above, we can construct a totally reduplicated form [$k^hæt^h$ $k^hæt^h$] cat cat in English, perhaps from a context like, ‘That is not a bobcat, it’s just a cat cat’ (for more detail on this form of reduplication, see Ghomeshi et al. 2004).

(134) *English total reduplication example*

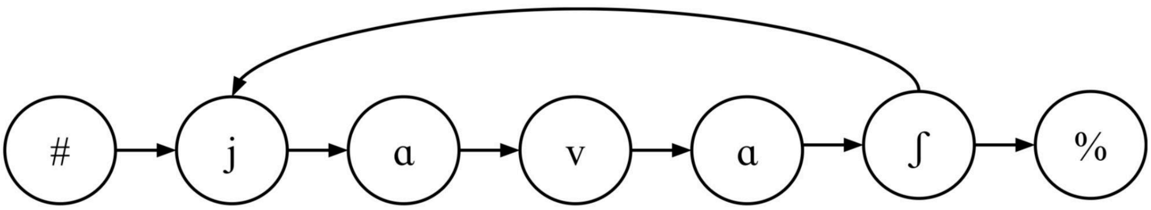


precedes k^h
k^h precedes æ
æ precedes t^h
t^h precedes k^h
t^h precedes %
output: [$k^hæt^h$ $k^hæt^h$]

When example (134) is spoken, the events {#, k^h, æ, t^h} are pronounced first (in that order). Then, rather than moving from /t/ to the % node, which would end the word, the edge from /t^h/ to /k^h/ is flowed through, and {#, k^h, æ, t^h} is pronounced again. The edge from /t^h/ to the % node will be traversed, and the word will end. Crucially, *the reduplicating edge is bounded by some kind of capacity such that it can be used only once*. That is, the loop between k and t cannot be infinitely traversed; [$k^hæt^h$] can be pronounced only twice. (See 2.1.3 for a brief discussion of capacities and flow networks in EFP.)

Turkish, like English, has total reduplication. This was also mentioned in the literature review (Chapter 2) but I will provide a summary here as well to introduce the graph representation provided below. In Turkish, total reduplication creates an adverbializing meaning. For example, the adjective *yavaş* [jɑvɑʃ] 'slow' can be fully reduplicated to form the adverb *yavaş yavaş* [jɑvɑʃ jɑvɑʃ] 'slowly'. Total reduplication in Turkish can be represented in the same manner as the English example above, with a reduplicative loop from the last phoneme of the word (in this case [ʃ]) back to the first first phoneme of the word (in this case [j]).

(135) *Turkish total reduplication example*[23]

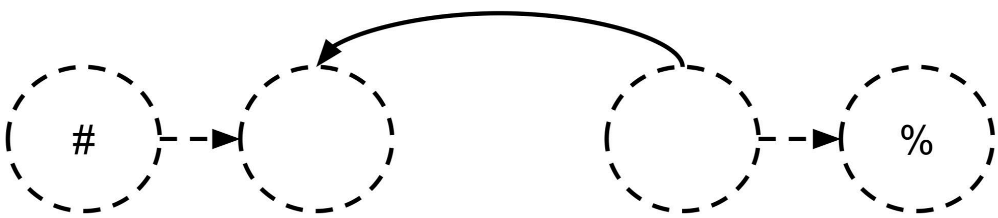


output: *yavaş yavaş* [jɑvɑʃ jɑvɑʃ] 'slowly'

Thus far, EFP graphs have been used to encode specific applications of total reduplication in English and Turkish on particular lexical items. However, EFP graphs can also be used to give a general structure of total reduplication. Broadly, total reduplication requires a reduplicating edge to be added from the last phoneme (the event preceding the sink %) back to the first phoneme (the event following the source #). Thus, the general structure of total reduplication is 'last event precedes first event'. Below in example (136) is what this looks like as an EFP directed graph structure.

[23] There is vowel harmony information which will be omitted from the graphs in this chapter and in Chapter 6, for ease of readability.

(136) *General structure of total reduplication 'last event precedes first event'*

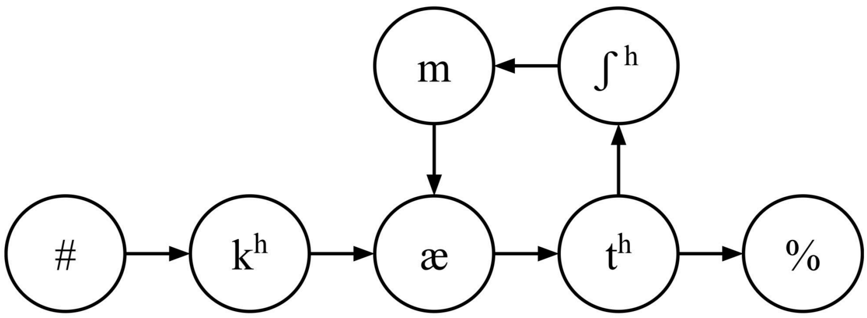


The dashed events and edges show the part of the structure that is matched to events and edges in a particular lexical item when the reduplication is applied to that item. The solid arrow at the top is the new edge which is added. The end result is that an additional edge is added from the end of the word back to the beginning, creating the total reduplication.

5.1.2 Representing partial reduplication

Next, let's turn to partial reduplication, beginning with an example: English shm-reduplication. English shm- reduplication is a partial suffixing reduplication in which nearly the whole base is copied except for the onset of the first syllable (if present), and the cluster shm- is added between the base and copied part. In this case, we need to insert two phonemes into the reduplicated form which are not present in the original form – [$ʃ^h$] and [m]. Example (137) below shows a graph of the English example [k^hæt^h $ʃ^h$mæt]:

(137) *English $ʃ^h$m- reduplication*

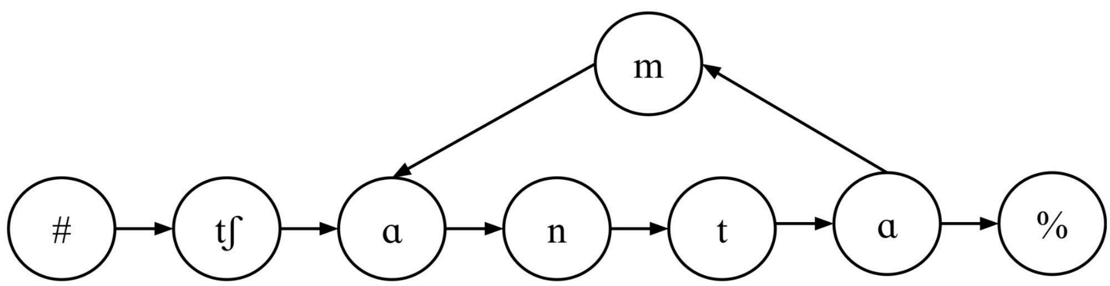


output: [k^hæt^h $ʃ^h$mæt^h]

In example (137), the reduplication cycle only includes some of the phonemes. This is because ʃm- reduplication is partial, and only the /æ/ and /t^h/ phonemes are be repeated in the pronunciation. When the graph in example (137) is spoken, the events {#, k^h, æ, t^h} will be pronounced first. Then, the reduplication cycle will be traversed {$ʃ^h$, m, æ, t^h}, in that order. Finally, the edge from /t^h/ to the % node will be traversed and the word will end, giving the output [k^hæt^h $ʃ^h$mæt^h]. For more information on this example, presented here briefly for illustrative purposes, see Raimy (2000:78).

As mentioned in the literature review, Turkish has a similar form of partial suffixing reduplication called m-reduplication. Almost all of the base is copied, with the initial onset (if present) being replaced with [m]. This type of reduplication has a generalizing meaning e.g., *çanta* [tʃɑntɑ] 'purse' can be reduplicated to form *çanta manta* [tʃɑntɑ mɑntɑ] 'purses and the like' (Göksel & Kerslake 2006). Unlike TER, this form of partial reduplication is productive, and the added consonant is invariably [m]. Turkish m-reduplication cannot be applied to words which start with [m] (Göksel & Kerslake 2006). This fact will be revisited later in my analysis of TER. Turkish m-reduplication can be represented using an EFP graph in much the same way as English $ʃ^h$m- reduplication.

(138) *Turkish m- reduplication*

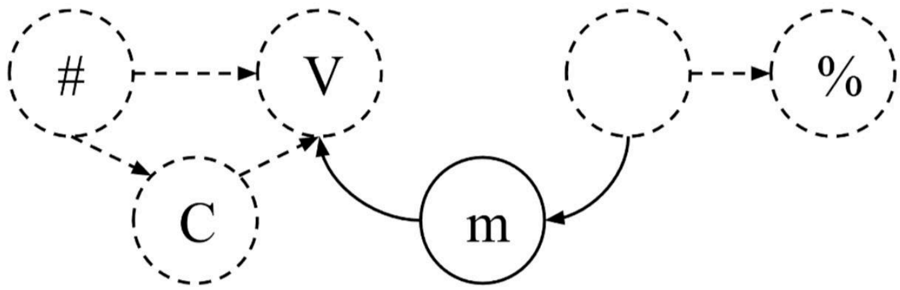


output: *çanta manta* [tʃɑntɑ mɑntɑ]

When pronounced, flow through the graph in (138) will begin at the source (#), followed by the events {tʃ ɑ n t ɑ} in that order. Then, the reduplicative loop with [m] will be flowed through back to the first [ɑ] vowel. Finally, {ɑ n t ɑ} will be flowed through once more, and traversal will finish at the sink (%). This leads to the correct output, *çanta manta* [tʃɑntɑ mɑntɑ] 'purses and the like'.

Finally, let's turn to the general structure of Turkish m-reduplication. The pattern of m-reduplication is that nearly the whole base is copied, except for the onset of the first syllable (if present), and an [m] is added before the reduplicated part. The structure of m-reuplication can be represented with 'm follows the final event and precedes the first event'. This prose description is represented in the following EFP graph:

(139) *Turkish m-reduplication 'm follows the final event and precedes the first event'*



As before, the dashed events and edges represent places where the reduplication process matches to the lexical item. Note that on the left side, there are two possible matches: the reduplicating edge attaches to the first vowel after the source (#), whether that vowel directly follows the source, or follows a consonant that follows the source. This allows for correct matching, whether or not there is an onset present on the first syllable. Secondly, note that in this case the reduplicating edge contains an [m]. This is because in every instance, application of m-reduplication results in an [m] being added between the base and the reduplicated part.

Now that we have covered the representation of both total reduplication and m-reduplication, we will turn to the main focus of this chapter: Turkish emphatic reduplication (TER).

5.2 How is the TER linker consonant selected?

TER is a partial prefixing reduplication which copies the onset of the first syllable (if present), along with the first mora of the base. In addition, when TER is applied, an additional ‘linker’ consonant is added. TER can be applied only to a limited set of adjectives and adverbs, and adds an emphasized meaning (Göksel & Kerslake 2006:90). See below in (42) for a table from section 2.5.1 of the literature review with examples:

(42) *Examples of Typical Partial Reduplication*

Bare Form	*Translation*	*Reduplicated Form*	*Translation*
güzel [gyzel]	'beautiful'	**güp**güzel [**gyp**gyzel]	'extremely beautiful'
mavi [mɑvi]	'blue'	**mas**mavi [**mɑs**mɑvi]	'very blue, stark blue'
eski [eski]	'old'	**ep**eski [**ep**eski]	'very old'
siyah [sijɑh]	'black'	**sim**siyah [**sim**sijɑh]	'very black, pitch black'
doğru [doːɾu]	'true, towards'	**dos**doğru [**dos**doːɾu]	'honest and aboveboard, dead ahead'
temiz [temiz]	'clean'	**ter**temiz [**ter**temiz]	'neat as a pin'

For consonant-initial words, the linker is selected from the set [p m s ɾ]. For vowel-initial words, the selected linker is always [p], though this may be slightly more complicated in reality (see the discussion below). In addition to the typical cases shown above, there are some words which can take an additional vowel or an additional vowel and consonant (Göksel & Kerslake 2006:90-91). See the table below, which contains two examples of such words also given in the literature review:

(140) *Additional linker examples*

Bare Form	*Translation*	*Reduplicated Form*	*Translation*
çıplak [tʃɯplɑk]	'naked'	**çırıl**çıplak [**tʃɯrɯl**tʃɯplɑk]	'stark naked'
sağlam [sɑːlɑm]	'durable, sturdy'	**sap(a)**sağlam [**sɑp(ɑ)**sɑːlɑm]	'as sound as a bell'

In terms of the process by which a linker consonant is chosen, there are several facts to consider which strongly indicate that the choice of linker consonant in TER is morpholexical and not the result of phonological processes. These facts are discussed at length in the literature review in section 2.5; see that section for much more detail on these issues than is given here.

First, TER is not productive as only a limited number of stems can undergo this reduplication process, perhaps several hundred. So, it is not outside the realm of possibility for the choice of linker on real TER forms to be the result of memorization for this finite set of forms. Second, there is a high degree of variation in the choice of linker for both real and nonce words (Taneri 1990, Wedel 1999, Kılıç & Bozşahin 2013, Demir 2017, Köylü 2021). Even vowel-initial words, which are said to always receive [p] as the linker consonant (e.g., Göksel & Kerslake 2006), do not seem to have a generalizable phonological process underlying them. In nonce word experiments, speakers have picked a wide array of linker consonants for vowel-initial tokens, not just [p] (see Demir 2017, Köylü 2021, Sofu & Altan 2008). There is even some limited evidence of a small number of speakers choosing [s] over [p] as a linker in real TER forms (see Taneri 1990 and Demir 2017). If the choice of linker consonant was truly governed by predictable rules or constraints, this degree of variation would not be found. Third, assuming lexicalization would also allow us to easily account for the highly exceptional forms like *çırılçıplak* [tʃɯ-ɾɯl-tʃɯplɑk] 'stark naked' without issue.

Finally, while many phonologists have been successful in using rules or constraints to capture generalizations about the choice of linker consonant, none have been successful in accounting for all of the data (see the literature review for a discussion of Taneri 1990 and Wedel 1999; Demir 2017 on Kelepir 1999; Yılmaz 2020 on Özçelik 2012; and Stachowski 2014 on Müller 2003). As a brief example, Wedel (1999) argues for a high ranking markedness constraint which blocks the linker from being identical to C2. While most attested TER forms do not take C2 as the linker, there are several exceptions. Phonological rules and constraints should not hold just most of the time; they should apply when structural criteria are met. Phonological processes such as dissimilation and identity avoidance fail to accurately capture the choice of linker for TER for all of the data because linker choice is simply not phonological. Rather, the choice of linker is morpholexical, where the linker consonant is stored with the base in the lexicon.

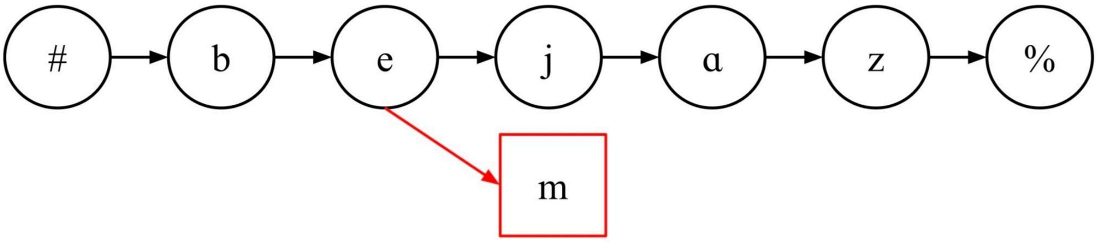
5.3 Closing the loop: using EFP to represent TER

While the choice of linker in TER is morpholexical, it is important to go beyond simply stating that the linker consonant is memorized and actually provide a structure for how this memorization occurs. In the following subsections, I will lay out, using the EFP framework, how linker consonants and TER can be represented using graphs.

5.3.1 The linker as a dangling event

In line with the fact that the linker is morpholexical, the choice of linker is stored in the lexicon with the base. See below for the EFP representation of the Turkish adjective *beyaz* [bejaz] 'white' in (141) below. This word is one of the limited number of Turkish adjectives which can undergo emphatic reduplication, and it is reduplicated as *bembeyaz* [be-m-bejaz] ‘completely white’.

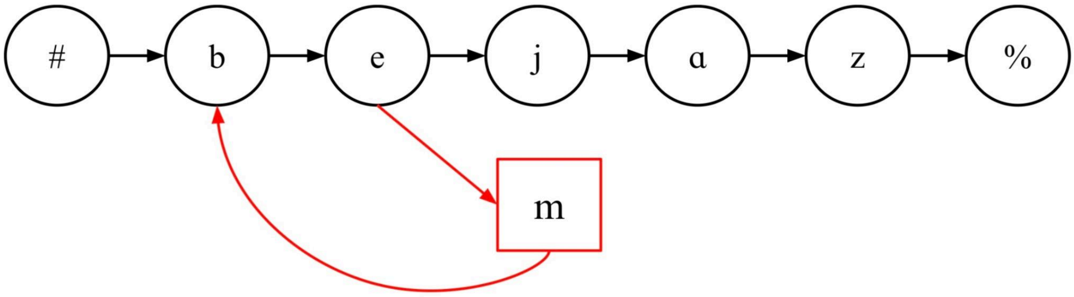
(141) *Unreduplicated Turkish word which can take TER*



The choice of linker consonant [m] is present in the lexical entry for the word *beyaz* [bejaz] as a ‘dangling event’ (red square) which follows the first vowel of the word. Note that this kind of dangling event is unique from the dangling events used in Chapter 4 to account for the vowels which sometimes appear in coda clusters. Those dangling events preceded another event but followed nothing (see subsection 4.4.2 for more details). As in Chapter 4, the red color and square shape of the event is not meant to indicate some kind of special ‘red square’ event in the phonological representation; they serve as a visual cue for the reader to recognize the dangling event. In terms of the actual phonological representation, in this case a dangling event is an event which precedes nothing and is not the sink (%).

When TER is applied to a word which normally takes it, a reduplicating edge is added so that the dangling event precedes the first event. See (142) below for a graph of the TER form *bembeyaz* [be-m-bejaz] ‘extremely white’. The event containing the linker [m] is connected back to the first event [b] using the added edge shown in red. Note that [m] is no longer dangling, as it now precedes [b].

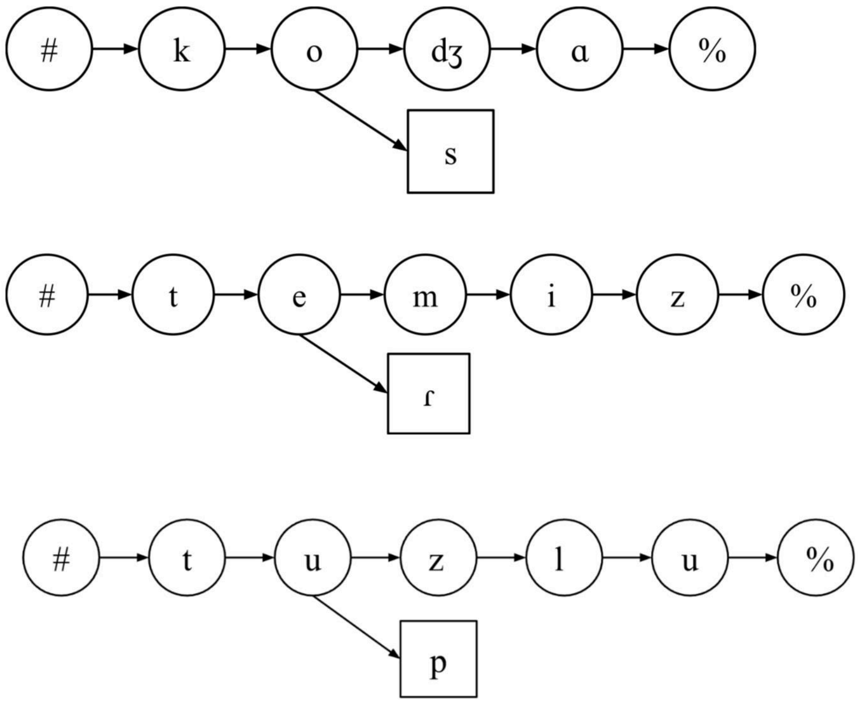
(142) *Emphatic reduplication*



In general, events which are left dangling are ignored in pronunciation. When a word which can take TER is pronounced without reduplication applied (as in 141 above), the dangling event is simply ignored and not pronounced. When a speaker applies TER to a word, they must attach an edge from the dangling event back to the first event of the word. Since the 'dangling event' now precedes another event, it is no longer dangling and is thus pronounced (as in 142 above).

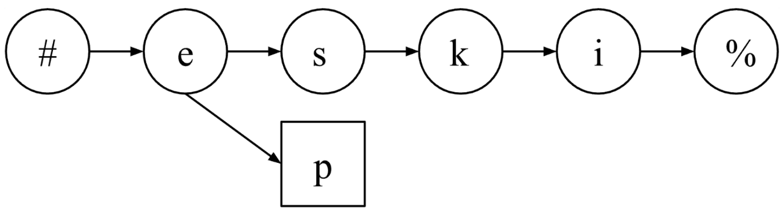
Other words which undergo TER can be represented in much the same way as *beyaz* [bejaz]. See (143) below for consonant-initial words which take each of the three other possible linkers {p s r}. These three words are *tuzlu* [tuzlu] 'salty' (TER: *tuptuzlu* [tu-p-tuzlu] 'extremely salty'), *koca* [kodʒɑ] 'huge' (TER: *koskoca* [ko-s-kodʒɑ] 'extremely huge'), and *temiz* [temiz] 'clean' (TER: *tertemiz* [te-ɾ-temiz] 'neat as a pin').

(143) *Consonant-initial bases which can undergo TER*



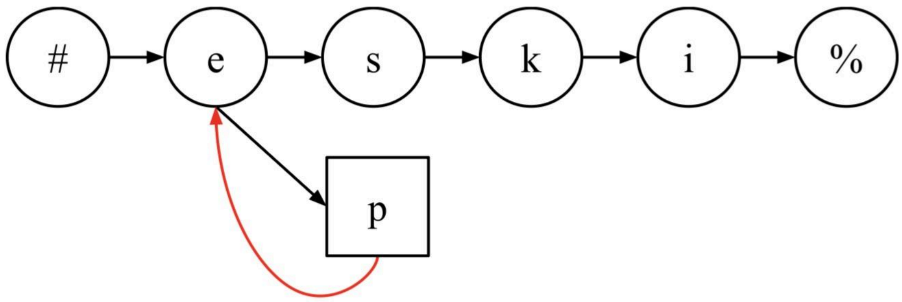
Vowel-initial words also have a linker which is stored on the word. The vast majority of the time this linker will be [p]. The linker consonant is morpholexical even on vowel-initial words because as mentioned above in section 5.2, there is some variation in what linker is chosen for vowel-initial forms. See (144) below for the example of *eski* [eski] ‘old’ which can be reduplicated to *epeski* [e-p-eski] ‘extremely old’.

(144) *Vowel-initial word which can undergo TER*



When TER is applied to a vowel-initial word, much the same process is applied as with consonant-initial forms: an edge is added such that the dangling event which contains the linker precedes the first event. The only minor difference is that for vowel initial words, the dangling event follows and precedes the same event (the first vowel) once TER is applied. This does not pose any issue, and the correct output is still attained.

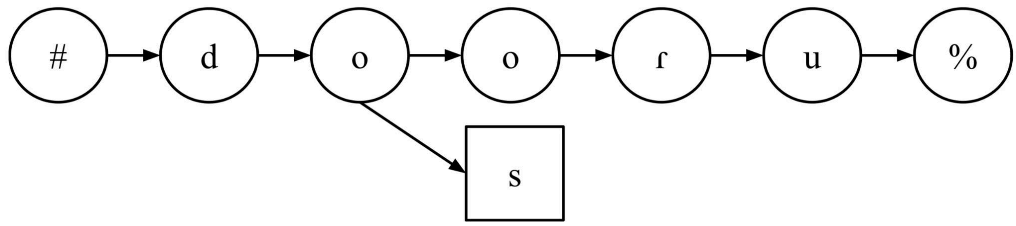
(145) *Vowel-initial word with TER applied*



There is one more small clarification to make about where the dangling event attaches to a lexical item. Recall that TER does not truly copy the first vowel, but actually copies just the first mora of the first vowel. When TER is applied to a word with a long first vowel, only the first mora is copied. For example, *doğru* [do:ɾu] 'true, towards' is reduplicated as *dosdoğru* [do-s-do:ɾu] 'honest and aboveboard, dead ahead'. Crucially, **doğsdoğru* *[do:-s-do:ɾu], where the copied vowel is long, is not correct. This does not pose an issue for the analysis of TER given here.

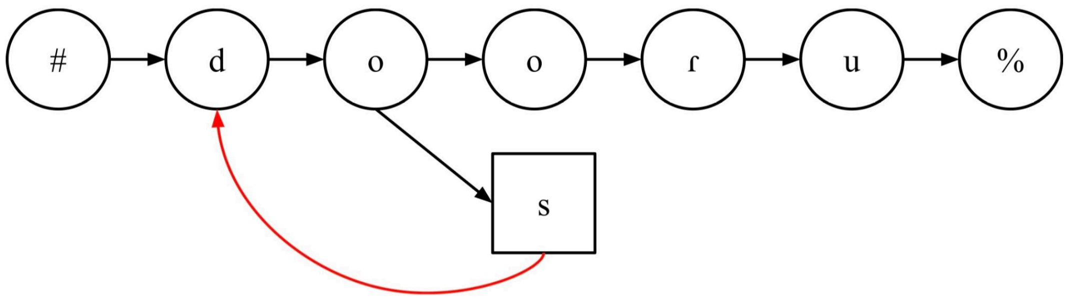
The length of a vowel can be represented in an EFP graph by how many timing events it takes up: a short vowel is one event, while a long vowel is two. Thus, a long vowel can be represented as two sequential identical vowel events. In the case of a word which takes TER with a long first vowel, the dangling event which contains the linker follows to the first vowel event. See (146) below for the EFP graph representation of *doğru* [do:ɾu].

(146) *Word with long first vowel that can undergo TER*



When TER is applied to *doğru* [do:ɾu], a new edge is added so that the dangling [s] precedes the first event [d] in the same manner described above for other words which undergo TER. See below for the representation with the reduplicating edge (red) added:

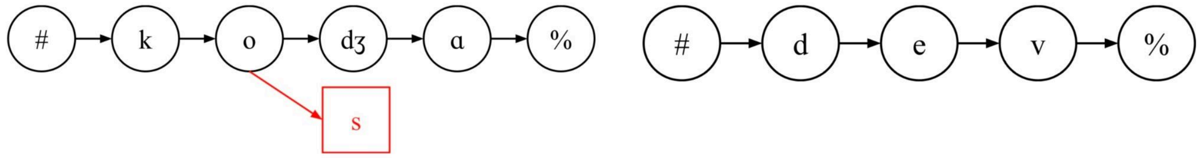
(147) *TER applied to word with long first vowel*



After the reduplicating edge is added, notice that the second [o] event is not included in the reduplication cycle. Thus, only the first mora of the long vowel is copied in the reduplication, leading to the correct output *dosdoğru* [do-s-do:ɾu].

One important benefit to including the dangling linker events on individual lexical items is that it provides a mechanism for marking which words can undergo TER. Recall that the set of words which undergo TER is unpredictable. While an adjective or adverb must be gradable to undergo TER (see Turgay & İskender 2021, Kaufman 2013), exactly which gradable adjective can undergo TER is arbitrary. For example, *koca* [kodʒɑ] 'huge' can undergo TER to become *koskoca* [ko-s-kodʒɑ] 'extremely huge' (shown above in 142). However, its synonym *dev* [dev] 'giant' does not undergo TER (Kaufman 2013). In terms of representation, a lexical item is marked for undergoing TER when it has a dangling event. Words which do not have a dangling event do not receive TER. See (148) below for an example with *koca* [kodʒɑ] 'huge' (left) and *dev* [dev] 'giant' (right).

(148) *Marking words for TER with a dangling event*

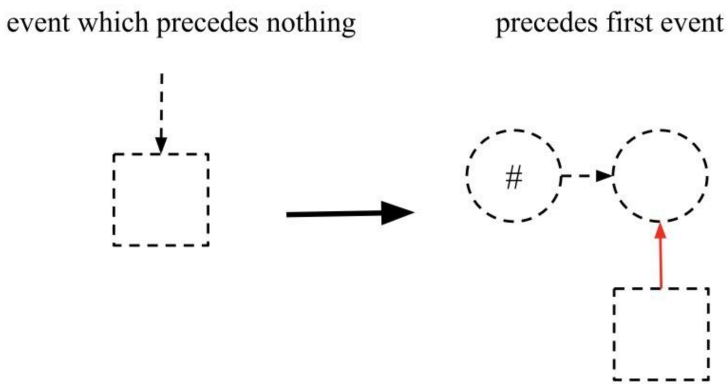


The word *koca* [kodʒɑ] 'huge' can undergo TER because it contains a dangling event (shown in red). Its synonym, *dev* [dev] 'giant' has no such dangling event and therefore cannot undergo TER.

5.3.2 Abstract representation of the TER process

While the words in Turkish which take emphatic reduplication have a linker consonant that is completely lexicalized, we can still investigate what aspects of TER *are* present in the morphology and phonology of this process. Since speakers are able to generalize emphatic reduplication to apply it to novel and nonce words (see Wedel 1999, Demir 2018, and Köylü 2021), some part of this process must be abstracted away from individual lexical items. In general, when speakers are applying TER, they are adding an edge from the dangling event back to the first event. Thus, the representation of TER includes the information 'an event which precedes nothing (dangling event) precedes the first event' which is shown as an EFP graph below:

(149) *TER Representation:* ‘an event which precedes nothing precedes the first event’



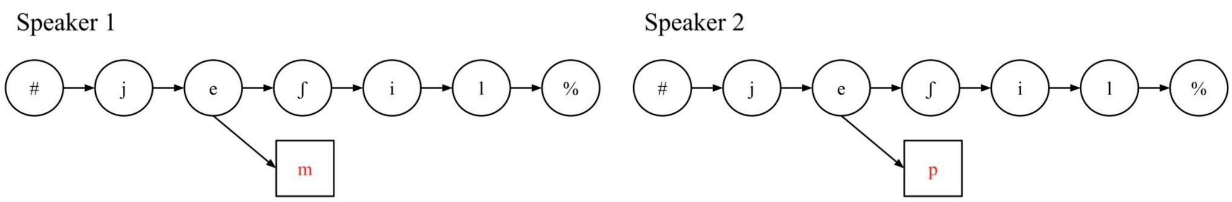
Note that for TER to apply, the dangling event must be one which precedes nothing. This blocks the process from applying to the previous kind of dangling event from Chapter 4 (the dangling high vowel event which occurs on words with ∅ ~ high vowel alternation). Those dangling events precede another event (but follow nothing), so the above structure cannot be matched appropriately, and the TER process will not interact with them.

Another important aspect of the general representation of TER shown in (149) is that it does not include any information about the linker consonant. The part that matches the dangling event (square) does not specify what linker is contained in the dangling event. The information about the linker consonant inside the dangling event is fully specified on lexical items which are eligible to undergo TER. When speakers are asked to apply TER to novel forms, this results in an ‘empty’ dangling event with no linker. Speakers then use various strategies to resolve the empty dangling event, leading to lots of variation in chosen linker. This will be discussed in more depth in section 5.4.

5.3.3 Representing variation in TER linker for real forms

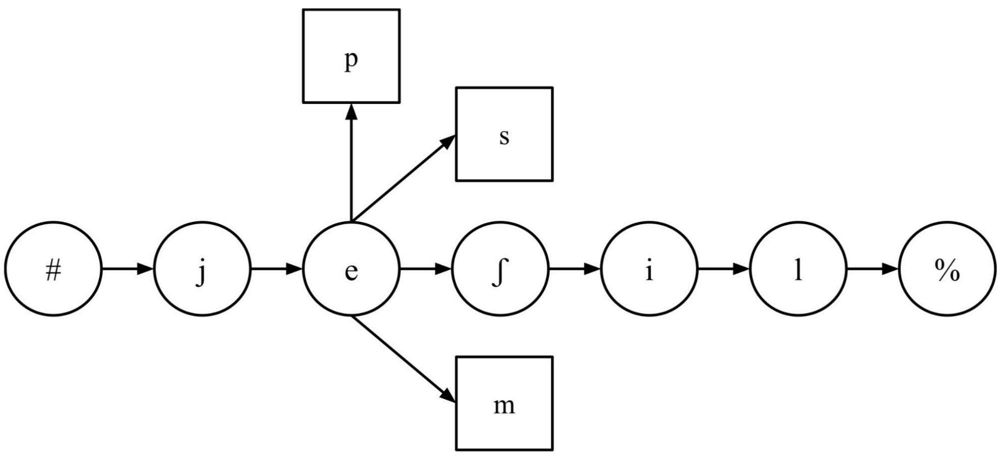
There is significant variation in the choice of linker for real attested TER forms. For example, the reduplicated form of *yeşil* [jeʃil] ‘green’ is reported to be *ye**m**yeşil* only (Wedel 1999, Göksel & Kerslake 2006), both *ye**m**yeşil* and *ye**p**yeşil* (Taneri 1990, Özçelik 2012, Yılmaz 2020), and *ye**m**yeşil*, *ye**p**yeşil*, and *ye**s**yeşil* (Köylü 2021). Where interspeaker variation exists, we can posit that different speakers have different linkers stored with a given lexical item. For example, if Speaker 1 accepts only *ye**m**yeşil* [je-m-jeʃil] and Speaker 2 accepts only *ye**p**yeşil* [je-p-jeʃil], these speakers have a dangling event which contains [m] and [p], respectively. These structures are shown below in (150).

(150) *Interspeaker variation in selected linker*



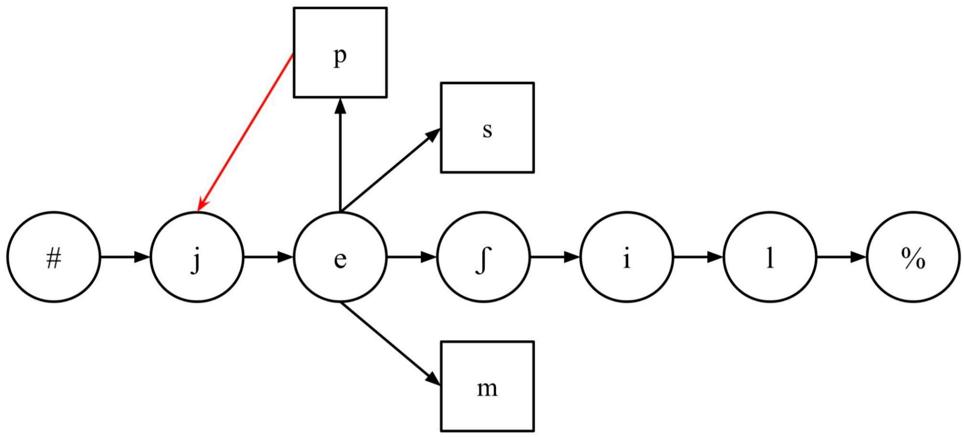
Intraspeaker variation, where a single speaker could produce any one of multiple linkers, can be represented using additional dangling events which all attach at the first vowel. This represents a choice, where the speaker may choose which of the dangling edges to attach the reduplicating edge to. The dangling event they choose will have the linker consonant they have selected to add to the word. See (151) below for a graph of the word *yeşil* [jeʃil] 'green' for a speaker who would use any of the three attested linkers [p m s].

(151) *Intraspeaker variation in linker*



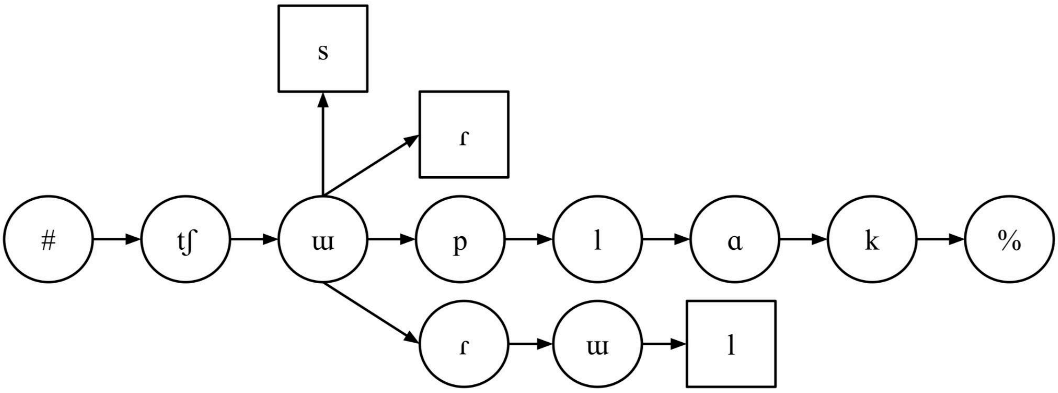
A speaker with a representation like that of (151) can choose whether to apply the reduplicating edge to any of the three dangling events [p], [m], or [s]. As an example, let us assume that the speaker chooses to use [p] as the linker. The appropriate representation is shown in (152) below with the reduplicating edge shown in red. When the speaker pronounces the word, the no-longer-dangling [p] is pronounced, while the still-dangling [s] and [m] events are ignored.

(152) *Graph of TER applied to word with multiple possible linkers*



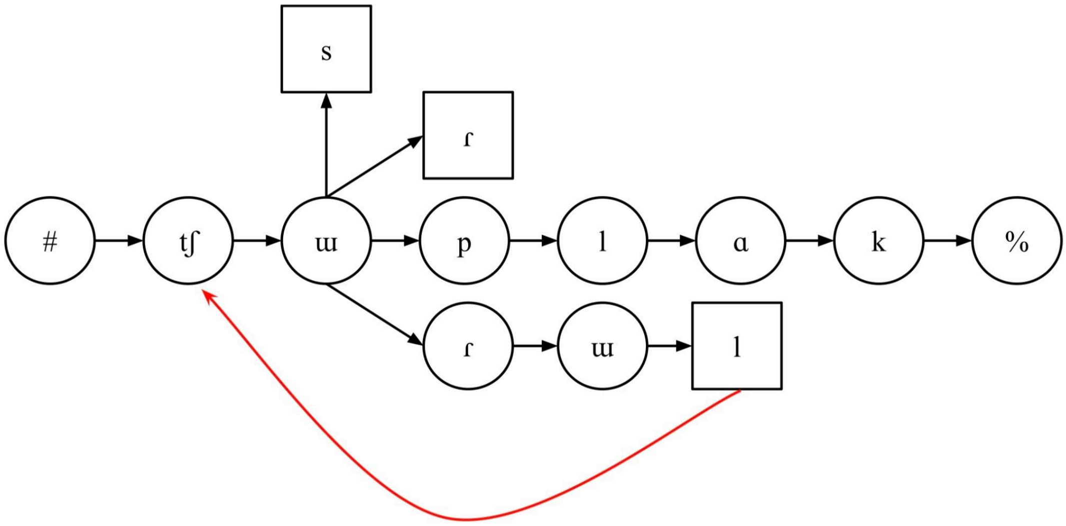
In addition to forms with multiple possible linkers, there are forms which can take more added material than a single linker consonant. For example, *çıplak* [tʃɯplɑk] 'naked' can be reduplicated as *çısçıplak* [tʃɯ-s-tʃɯplɑk], *çırçıplak* [tʃɯ-ɾ-tʃɯplɑk], and also as *çırılçıplak* [tʃɯ-ɾɯl-tʃɯplɑk], with an extra added vowel and consonant. The representation for this form is shown in (153) below. Specifically, this is a representation for a speaker who would accept *çısçıplak* [tʃɯ-s-tʃɯplɑk], *çırçıplak* [tʃɯ-ɾ-tʃɯplɑk], or *çırılçıplak* [tʃɯ-ɾɯl-tʃɯplɑk]. The possibilities of [s] and [ɾ] as linkers are represented with dangling events in the same manner as the previous example in (151). To capture the possible form *çırılçıplak* [tʃɯ-ɾɯl-tʃɯplɑk], the additional consonant and vowel are simply added before the final linker [l].

(153) *TER form which can take extra added material*



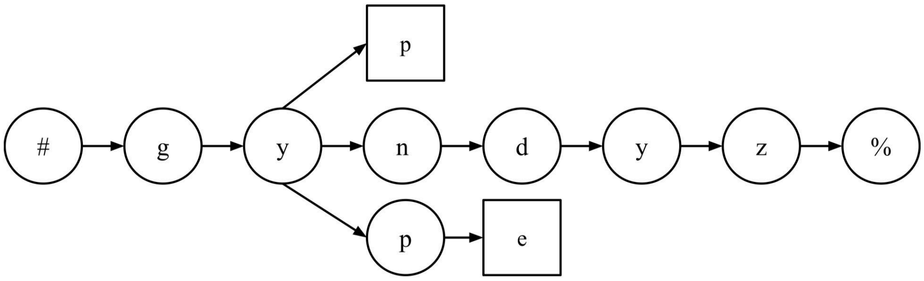
If a speaker chooses to pronounce *çırılçıplak* [tʃɯ-ɾɯl-tʃɯplɑk], they will attach the reduplicating edge from the [l] to the first event [tʃ], shown below in red.

(154) *TER applied with extra added material*



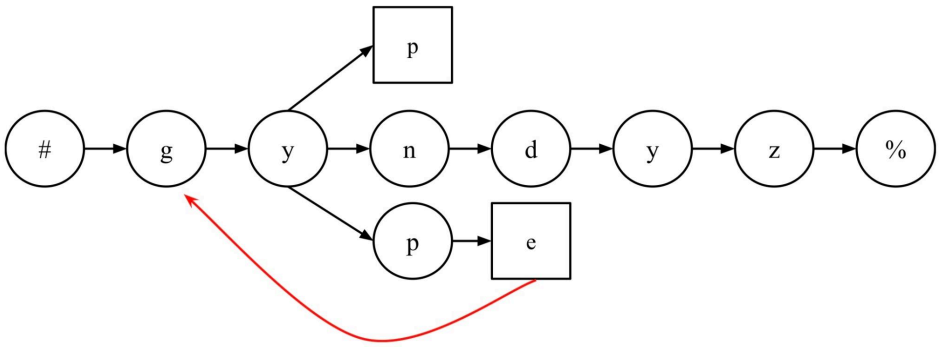
This type of structure can be used to represent all other forms which take extra added material. For instance, some words can take an optional extra vowel when TER is applied. When *gündüz* [gyndyz] 'by day' *undergoes TER*, it can take the forms *güpgündüz* [gy-p-gyndyz] or *güpegündüz* [gy-pe-gyndyz] with an additional added [e].

(155) *TER form with extra added vowel*



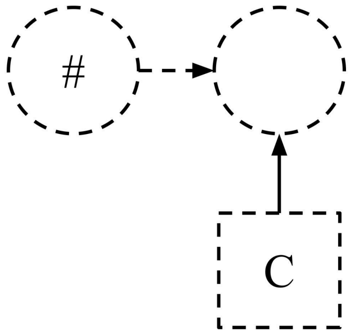
When TER is applied, the reduplicating edge connects the dangling event [e] to the first event [g]. The resulting graph is shown below in (156).

(156) *TER form with extra added vowel*



Note that the lack of specification of linker in the general structure of TER (given in 149) is crucial for applying TER to these forms with extra added material. The dangling event which the reduplicating edge attaches to is not limited to [p m s ɾ]. As shown above, the dangling event could be another consonant (such as [l] in 153) or even a vowel (such as [e] in 155). Any attempt to specify what kind of phoneme should be matched in the dangling node would cause the reduplication to not apply properly in some of these cases with extra added material. For example, we could posit that the dangling event is specified to contain a general consonant, as in (157) below.

(157) *Match a dangling event with a consonant*

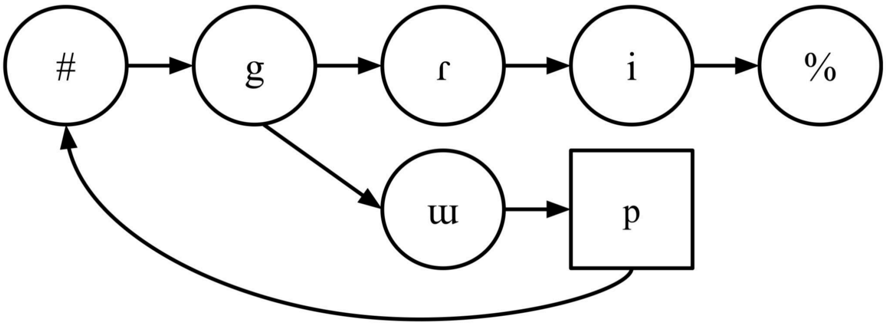


However, the dangling consonant event in this structure would fail to match to the dangling [e] event in (155), and there would be no way to form *güpegündüz* [gy-pe-gyndyz] as shown in (156). Thus, leaving the dangling linker event completely unspecified in the abstract structure of TER is critical for correct application in all cases.

5.3.4 Returning to gıpgri [gɯ-p-gɾi]

In Chapter 3, it was mentioned that the form *gıpgri* [gɯ-p-gɯɾi] ‘extremely gray’ may pose an issue for the argument that vowels which occur in onset clusters in Turkish are intrusive vowels. The reason being that intrusive vowels should not interact with the phonology; thus, it would seem very odd that the intrusive vowel in *gri* [gɯɾi] would be reduplicated when TER is applied. I argue that the vowel in this case actually is not part of the reduplicated part of the base. Below is the EFP representation of this form with TER applied:

(158) *EFP representation of gɪpgri [gɯ-p-gurɾi] 'extremely gray'*

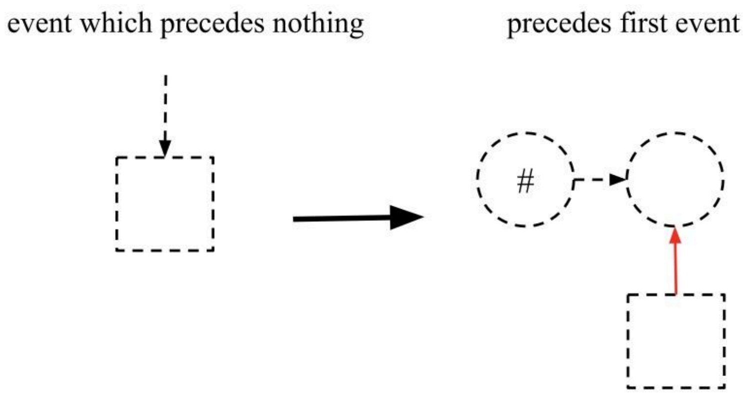


It has already been demonstrated that it is necessary in some cases to include vowels in preceding the dangling consonant (see 153 and 155 above). In this case, the reduplicated part of the base is just the initial consonant [g] with the [ɯ] preceding the dangling edge, parallel with the phonemes of the word, as with previous examples that include extra vowels (153 and 155). Thus, this word need not pose any issue with the idea that the vowels inserted into onset clusters are intrusive. The vowel in the onset cluster does not need to be present in the phonology for this form to have the correct surface form, given the representations laid out here.

5.4 Trying to fill in the blanks: applying TER to novel forms

Only representations for real, attested TER forms have been demonstrated so far. When asked to apply emphatic reduplication to novel words and nonce words, speakers used a diverse array of strategies for filling in the linker consonant. First, consider, given the structure posited here, what is happening when a speaker applies TER to a novel word or a nonce form. Recall that this is all that is present in the phonology of TER:

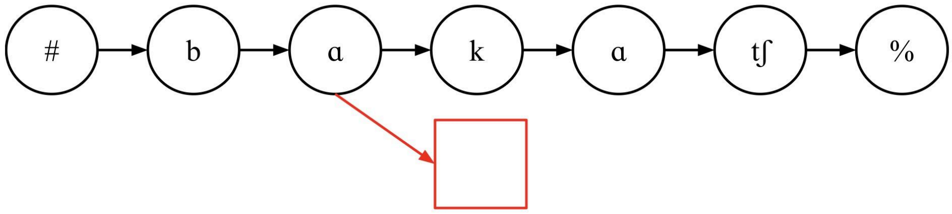
(149) *TER Representation: 'an event which precedes nothing precedes the first event'*



As discussed in 5.3.3, no specification for the selection of linker is present in this structure. The linker is present on those certain lexical items which undergo TER as a dangling event. When TER is applied to a nonce or novel form, there is no linker specified, meaning that applying TER to nonce or novel forms is a different kind of task than applying it to a real word which normally can undergo TER. This section lays out what is happening, structurally, for each of these two new cases (nonce forms and novel words).

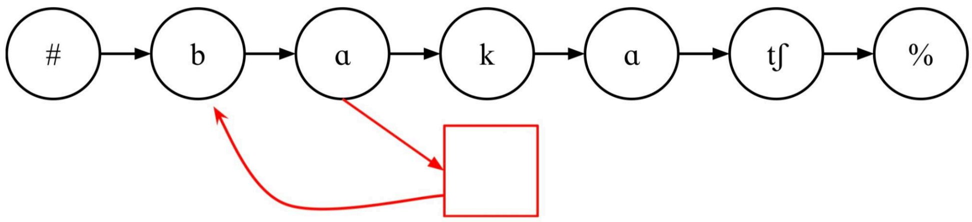
In an experimental task with nonce words, a speaker must construct a new entry in their lexicon for each nonce word. Since the task instructs participants that these nonce words are eligible to be reduplicated, they must be added to the lexicon as a new entry with a dangling event, because the presence of a dangling event is what marks a word as being eligible to undergo TER. However, since these are completely new entries into the lexicon, and the phonemes of the word give no indication as to the identity of the linker, the dangling event will not be filled with any feature information. See below for an example graph structure of the Turkish nonce *bakaç* [bakatʃ], which was used in Demir (2017).

(159) *Lexical entry for nonce word bakaç [bakatʃ]*



When a speaker applies TER to this nonce, they will add an edge such that the dangling event precedes the first event, as per the general structure of TER. See the following graph for TER applied to the example above:

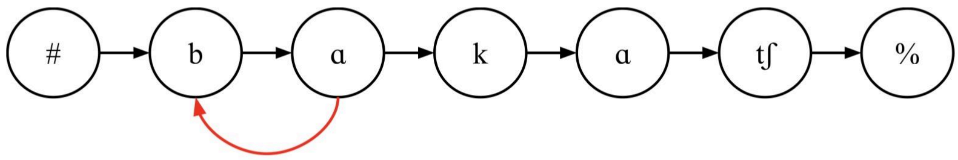
(160) *Nonce bakaç [bakatʃ] with TER applied*



The issue with the structure shown in (160) is that there is no feature information inside the linker event (red square). It is simply empty. An empty event cannot be pronounced, so this needs to be resolved in some way. Since this nonce word was used in a real experimental study, there is some information available on what speakers might do in this situation. Demir (2017) reported that 40 speakers chose *basbakaç* [ba-s-bakatʃ], 36 chose *bambakaç* [ba-m-bakatʃ], 7 chose *bapbakaç* [ba-p-bakatʃ], and 6 chose *barbakaç* [ba-ɾ-bakatʃ]. Note that in attested TER forms, [p] is never selected as a linker for [b] initial words, and 7 speakers chose that option for this nonce word.

In addition to the above, a further 26 speakers did not choose a linker from [p m s ɾ] when applying TER to *bakaç* [bakatʃ]. This category included omitting the linker entirely, using C2 as the linker (in this case [k]), or using a seemingly random phoneme. Unfortunately, Demir did not give the details as to which one of these three options those 26 speakers used in this specific case. While the variety of responses may seem at first odd, they make perfect sense when viewed from the lens of attempting to resolve an invalid, empty event. For instance, one way of resolving an empty event is to simply delete it, which leads to no linker added in the surface form. The graph below shows what this looks like for the nonce *bakaç* [bakatʃ]:

(161) *Nonce bakaç [bakatʃ] with empty event deleted*

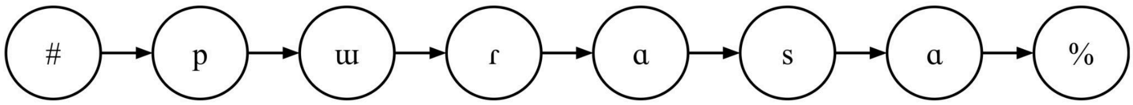


Since the empty event is no longer present in the graph (as it was in 160), the speaker has successfully resolved the issue, and the form is pronounceable. In this case, the surface form will be *babakaç* [babakatʃ].

Much more detail about exactly what strategies speakers might be employing to resolve the empty linker event will be provided below. However, suffice it to say that this empty event explains the variation in linker choice found in the previous nonce word studies of TER (e.g., Demir 2017, Köylü 2021).

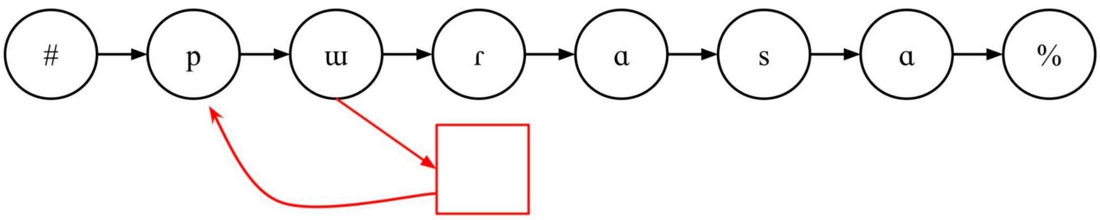
A similar issue arises with an empty event when speakers are asked to reduplicate real words which do not typically undergo TER (as in Wedel 1999 and Kılıç & Bozşahin 2013). As mentioned in 5.3.1 (see figure 148 and accompanying discussion), Turkish words which do not undergo TER lack a dangling event. Kılıç & Bozşahin (2013) asked speakers to apply TER to the Turkish word *pırasa* [pɯɾɑsɑ] 'leek'. Shown below in (162) is the lexical entry for this word. Note that since this form is not part of the limited set of words which can normally undergo TER, it lacks a dangling event.

(162) *Lexical entry for pırasa [pɯɾasa] 'leek'*



When speakers applied TER to this lexical item, they use the structure shown in (149), which specifies that a dangling event must precede the first event. Since no dangling event is present in (162), a dangling event is created and a reduplicating edge is added between the new dangling edge and the first event. The resulting structure (shown in 163 below) now conforms to the general structure of TER, and is analogous to a nonce word with TER applied as shown in (160).

(163) *Turkish word pırasa [pɯɾasa] 'leek' with TER applied*



As with the nonce word example in (160), speakers are left with an empty linker event (red square in 163 above). This empty event must be filled in. Since this word was included in a study, there is data available about how speakers go about filling in the empty event in this case. Kılıç & Bozşahin (2013) reported that 17 speakers chose *pıspırasa* [pɯ-s-pɯɾɑsɑ], 14 chose *pıppırasa* [pɯ-p-pɯɾɑsɑ], 10 chose *pırpırasa* [pɯ-ɾ-pɯɾɑsɑ], and 9 chose *pımpırasa* [pɯ-m-pɯɾɑsɑ]. Unlike Demir (2017), Kılıç & Bozşahin (2013) did not observe any speakers not using a linker from the set [p m s ɾ]. Like Demir (2017), however, Kılıç & Bozşahin (2013) observed speakers selecting a linker for a nonce form which is not attested in real forms. For real TER forms, C1 is never chosen as the linker, but nearly half of the speakers in this study chose the linker [p] for a [p] initial form in this case.

It may seem strange that speakers would apply to novel and nonce forms linkers which would not be attested on a similar real form. However, this finding makes sense when one considers applying TER to nonce and novel forms as a *fundamentally different task* from applying it to words which normally undergo TER. When applying TER to a word which normally undergo it, the linker is fully specified in a dangling event attached to the lexical entry for the word. TER application involves simply adding an edge from the dangling event to the first event. When applying TER to novel and nonce forms, on the other hand, a speaker must add an empty dangling event which did not previously exist and add the reduplicating edge from that new dangling event to the first event. Then, since an empty event is not pronounceable, the speaker must find a way to resolve it, either by filling it in with features, or deleting it completely.

5.4.1 Foraging and probability matching

While speakers do vary widely in their choice of linker on novel and nonce forms, their choices do not seem to be completely random– there do seem to be some 'fuzzy' patterns in the data. These patterns are 'fuzzy' in that they have exceptions.

Firstly, in the broadest sense of a pattern, speakers do tend to stick to the linkers found in attested forms when applying TER to novel words or nonce forms. Some speakers assume, from experiences with words that do receive TER, that the empty linker event must be filled in with [p m s ɾ], or even just [p m s], since [ɾ] appears so rarely as a linker consonant on attested forms. For example, Köylü (2021) found that all 14 speakers surveyed used linker consonants from the set [p m s] when reduplicating nonce words, despite the task being free response. Similarly, Wedel (1999) found that participants always picked [p m s] for the linker consonant when applying TER to novel Turkish words. However, not all speakers do this- both Demir (2017) and Sofu (2005) have observed speakers using linkers not from the set [p m s ɾ] when applying TER to nonce words.

Beyond that, speakers may even have intuitions about the statistical likelihood of a given linker consonant being near to certain phonemes in the base. This leads to these 'fuzzy' patterns, where most (but not all) speakers will pick one or two linker consonants for a given novel or nonce word. In general, there seems to be a pressure for the linker to be less similar to consonants in the base, especially C1 and C2. Phonologists have argued that this is a case of dissimilation (see Kelepir 2001).

In Wedel's (1999) study with novel words, for instance, speakers picked the same one or two consonants from the set [p m s] for most of the words, and the linker consonant selected tended not to match the first two consonants of the base. However, some words were reduplicated with a mix of consonants (for example, *tecrübeli* [tedʒɾybeli] 'experienced' was reduplicated with [p] by two speakers, [s] by two speakers, and [m] by one speaker). In addition, a small number of responses contained a linker consonant that was identical to either the first or second consonant of the base.

As a specific example, there seems to be a dispreference for selecting [s] as a linker on words which begin with ş [ʃ] and [z] (and [s] as well, but this can be covered by a general dispreference for the linker to be identical to C1, which is discussed in the next section). For real attested forms, there are no words which start with ş [ʃ] or [z] that take [s] as a linker (see Taneri 1990 and Yılmaz 2020, for example). In novel and nonce word studies, [s] seems to be dispreferred as a linker for words which start with ş [ʃ] or [z]. Kılıç & Bozşahin (2013) report that for the novel words *şerit* [ʃeɾit] 'band, strip' and *zarf* [zɑɾf] 'envelope', not a single speaker out of the 30 included in the experiment selected [s] as a linker for either. Demir (2017) found that out of 106 speakers, only 3 chose [s] as a linker for the nonce form *şipen* [ʃipen] and only 1 chose [s] as a linker for the nonce form *zugan* [zugɑn]. Köylü (2021) found that out of 14 speakers, 4 chose [s] as a linker for the nonce form *zımı* [zɯmɯ], while 8 chose [s] as a linker for the nonce form *zıpı* [zɯpɯ].

I argue that the dispreference for [s] as a linker on novel or nonce words which start with ş [ʃ] or [z] is not the result of some kind of dissimilation process. These 'fuzzy' patterns, in general, are better explained by speakers resorting to non-phonological means of resolving the empty linker event in novel forms.

One such non-phonological strategy may be related to mental foraging. Foraging behavior has been studied in a number of animals in terms of literal food gathering (for example, see Anselme & Güntürkün 2019). However, foraging behavior has also been observed on a mental level in humans during lexical access tasks (Lundin et al. 2023). Essentially, it is argued that speakers form mental 'patches' of grouped words, which they access when completing such tasks. It is possible that Turkish speakers form patches of words which have different kinds of dangling linker consonants (e.g., a patch for words which receive [p] as linker, word receive [s], etc) and then consult these patches when attempting to fill in the empty linker event in a novel form. The makeup of the mental patch may inform the choice of linker. For example, the lack of words which start with ş [ʃ] or [z] in the patch of words with an [s] linker, and may then disprefer [s] as a linker on novel forms which start with sounds.

There is some evidence which suggests that when speakers are faced with inconsistent linguistic input, they do not always regularize. Rather, speakers may engage in probability matching, where their productions match the probability of inconsistencies in the input (Wonnacott & Newport 2005). Thus, if a speaker knows that words which start with a certain consonant typically, but do not always, take a certain linker, most speakers may pick the typical consonant, while others reflect the inconsistency in the input by picking a less common linker. A combination of mental foraging and probability matching could explain these 'fuzzy patterns' where most speakers, but not all, Turkish speakers will pick a certain linker for a given novel form.

To my knowledge, there is only one study on TER which has directly examined strategies for linker selection related to mental foraging or pattern matching. Kılıç & Bozşahin (2013) found that consonant clusters formed by the linker and first consonant of the base in novel forms seemed to be preferred when they were *less* likely to occur in Turkish words in general (see section 2.5.8 of the literature review for more information on this study). This finding seems to tie linker selection with some kind of reverse probability matching to the broader lexicon.

This section is very high level and speculative, and that is by design. These strategies are notably not phonological ones, and this dissertation is ostensibly focused on phonological representations. My purpose here is not to provide a definitive account of linker selection by these means, but rather to gesture to possible (potentially more effective) alternatives to a phonological approach.

5.5 Dispreferred or disallowed structures

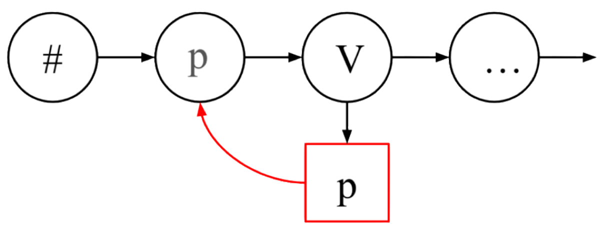
Aside from dissimilation, phonologists have argued that the choice of linker is a product of identity avoidance (e.g., Wedel 1999, Köylü 2021, Tang & Akkuş 2023). Evidence for this that speakers have a general dispreference for selecting a linker which is identical to C1 or C2. I argue, however, that this 'identity avoidance' is better explained using structural constraints on the graph, rather than some kind of dissimilation process.

5.5.1 C1 as a 'bad' linker

Despite significant variation in linker choice, a linker consonant is never identical to C1 of the base in real TER forms. For example, the TER form of *perişan* [peɾiʃɑn] 'miserable' has been most frequently attested as perperişan [pe-ɾ-peɾiʃɑn] (Göksel & Kerslake 2006, Wedel 1999, Demir 2017, Yılmaz 2020, Stachowski 2014), with some sources attesting *pesperişan* [pe-s-peɾiʃɑn] (Demir 2017, Yılmaz 2020) and *pemperişan* [pe-m-peɾiʃɑn] (Demir 2017). Using a linker which is identical to C1 to form **pepperişan* *[pe-m-peɾiʃɑn], to my knowledge, has never been attested in any source. Using a linker identical to C1 in novel and nonce word studies is reported to be dispreferred (Wedel 1999, Kılıç & Bozşahin 2013, Demir 2017, Köylü 2021). However, there are cases where C1 is chosen as the linker for novel and nonce forms. For instance, the aforementioned form *pıppırasa* [pɯ-p-pɯɾɑsɑ] from Kılıç & Bozşahin 2013 (chosen by 14 speakers) demonstrates this fact.

Rather than being a product of some kind of dissimilation or identity avoidance, this fact has to do with maintaining a valid graph structure of the reduplication itself. When a consonant identical to C1 is chosen as the linker, there is an equivalent graph structure which does not conform to the general structure of TER. See (164) below for an example of applying [p] as the linker on a word which starts with [p].

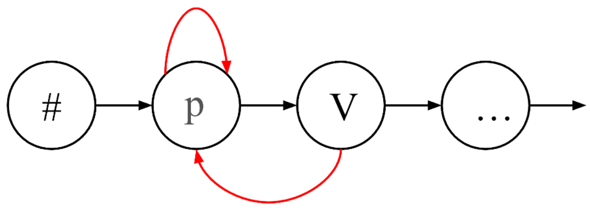
(164) *Graph of [p] initial word with [p] linker*



output: pV-p-pV…

Since C1 and the linker are both [p], they can be collapsed into a single event, as shown below in (165):

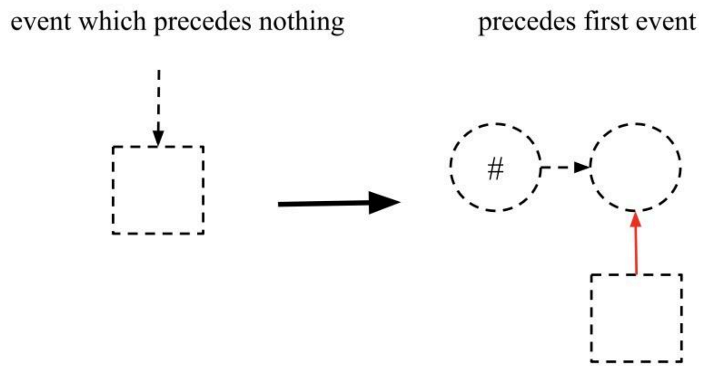
(165) *Alternate equivalent graph of [p] initial word with [p] linker*



output: pV-p-pV…

The structure in (165) has the same output as that of (164). Flow begins at the source and continues to the first vowel (V). Then, the bottom reduplicative edge is flowed through. At this point, flow is back at the [p] event. The top loop is then flowed through, so [p] is output again. Finally, flow moves back through the first vowel (V) and onto the rest of the word. This yields the output *pV-p-pV*, which is the identical to that of (164). While the correct output is attained, notice that the structure in (165) does not match the general structure of TER. As a reminder, the general structure of TER is ‘an event which precedes nothing precedes the first event’, as shown below:

(149) *TER Representation: 'an event which precedes nothing precedes the first event'*



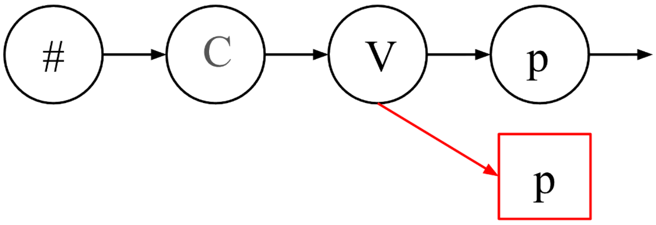
There are a couple ways in which the structure shown in (165) diverges from the general structure for TER given in (149). In the graph in (165), there are two events (aside from the source) which precede the first event. The first vowel (V) precedes the first event. In addition, the first event [p] also precedes itself. This is the first issue: the first event in (165) is preceded by two events, while the general structure of TER only has one event preceding the first event (again, aside from the source). The second issue is that neither the [p] nor the V event are dangling. This structural mismatch between the equivalent graph representation and the general structure for TER is what leads to C1 being banned as a linker (in real TER forms) or highly dispreferred (in novel and nonce word studies).

5.5.2 C2 as a 'bad' linker

In general, speakers using a linker which is identical to C2 on real TER forms is rare (though there are a few attested forms where this is the case). It has also been reported in some novel and nonce word studies that speakers disprefer using C2 as a linker (Wedel 1999, Kılıç & Bozşahin 2013, Demir 2017, Köylü 2021), though it does not seem to be as rare as in existing TER forms.

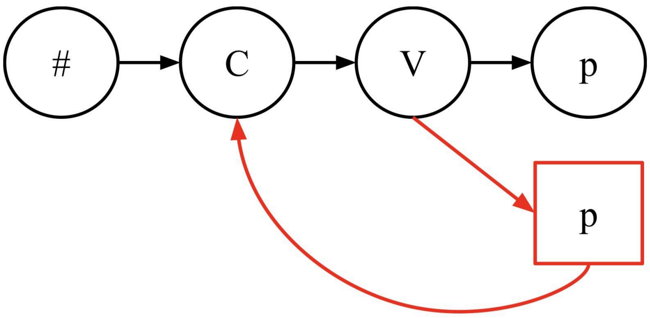
This dispreference for using C2 as a linker can also be explained using the structure of the EFP graph representations. The graph in (97) shows an example word in which C2 and the linker (red) are both [p].

(166) *Graph of word with identical C2 and linker*



Notice that in the graph above there are two identical [p] events in parallel to each other. That is, there are two identical events which both follow the same event (V). This structure with two parallel, identical events is dispreferred. The following example shows the structure in (166) after TER is applied and an edge is added between the dangling [p] event and the first event:

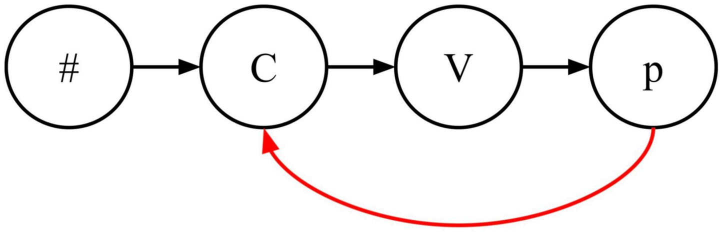
(167) *Graph of TER applied to word with identical C2 and linker*



output: CV-p-CVp

The two parallel [p] events in the above graph can be combined to form an equivalent graph structure with the same output, shown in (168):

(168) *Alternate equivalent representation of word with identical C2 and linker*



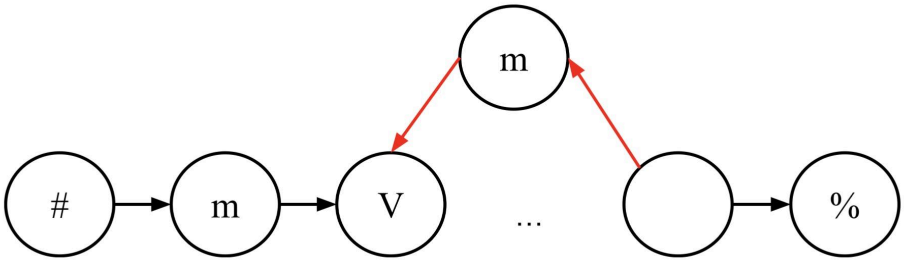
output: CVp-CVp

In (168), the first three events are reduplicated, rather than just the first two. Note that this graph does not match the structural description of TER which is shown in (149). The structural description of TER consists of 'an event which precedes nothing precedes the first event', and the structure in (168) lacks a dangling event entirely.

This dispreference for parallel identical events is not isolated to TER. Turkish m-reduplication is a partial suffixing reduplication in which the initial onset is removed (if present) and replaced with [m] in the copied part. See 4.1.2 for background on this form of reduplication and how it can be represented using EFP. Recall that Turkish m-reduplication is blocked on words which start with [m].

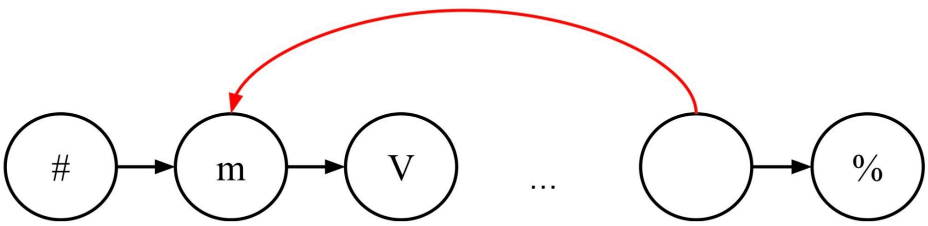
Shown below is an example graph of m-reduplication being ungrammatically applied to an [m] initial word:

(169) *Graph of m-reduplication applied on an [m] initial word*



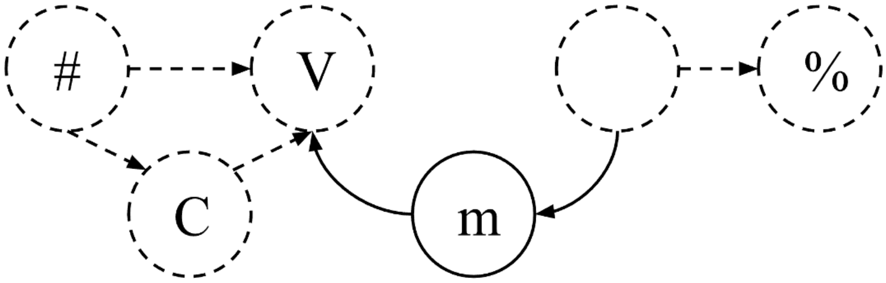
As in (167), there are two identical events in (169) which are parallel to each other. In this graph, there are two [m] events in parallel which both precede the first vowel (V). As before, these two parallel identical events can be collapsed together, to form the structure given below:

(170) *Equivalent graph of m-reduplication applied on an [m] initial word*



The structure in (170) does not match the general structure of Turkish m-reduplication which was given in 5.1.2, also pictured below in (139). The graph in (170) has a reduplicating edge from the last event to the first onset, unlike the general structure in which the reduplicating edge attaches to the first vowel. Secondly, there is no [m] along the reduplicating edge in (170).

(139) *General representation of Turkish m-reduplication*



Furthermore, the reduplicating edge in (170) goes from the last event to the first event, which matches the structure for total reduplication, which was discussed in 4.1.1, which likely increases dispreference for this structure. Total reduplication exists in Turkish and has a completely different meaning from m-reduplication, which is potentially another factor which blocks m-reduplication from applying to [m] initial words.

5.6 Conclusion

The EFP framework was used to provide a graph-based account of TER. For Turkish words that take TER (e.g., *beyaz* [bejɑz] white), the choice of linker consonant is completely lexicalized, being memorized on each root that takes TER. However, speakers do have a generalized structure including adding a precedence relationship from a dangling event which precedes nothing to the first event. This general structure notably does not include any information about the linker consonant contained in the dangling event. When asked to reduplicate novel and nonce forms, speakers use various strategies to resolve the empty linker consonant, including using foraging and probability matching based on existing forms they know, deleting the linker consonant altogether, and reduplicating the second consonant of the base. Each of these cases was discussed, and where appropriate, an EFP graph representation was provided.

6 Plosive Laryngeal Alternation

This section discusses several important topics related to plosive laryngeal alternation. Some plosives at the ends of words alternate in laryngeal setting when resyllabified into the onset of a vowel initial suffix. However, there are some final plosives which do not alternate. See the table below (taken from subsection 2.6.1 of the literature review) for examples of each kind of final plosive in roots.

(59) *Types of laryngeal setting in Turkish plosives*

	Phone	*Bare Form*	*Translation*	*Suffixed Form*	*Translation*
a.	p	küp [kyp]	‘cube’	küpü [ky.p-ym]	‘my cube’
	t	at [ɑt]	‘horse’	atım [ɑ.t-ɯm]	‘my horse’
	ç [tʃ]	saç [sɑtʃ]	‘hair’	saçım [sɑ.tʃ-ɯm]	‘my hair’
	k	bank [bɑnk]	‘bench’	bankım [bɑn.k-ɯm]	‘my bench’
b.	b	tab [tɑb]	‘print’	tabım [tɑ.bɯm]	‘my print’
	d	ad [ɑd]	‘name’	adım [ɑ.d-ɯm]	‘my name’
	c [dʒ]	sac [sɑdʒ]	‘sheet metal’	sacım [sɑ.dʒ-ɯm]	‘my sheet metal’
	g	ring [ɾing]	‘boxing ring’	ringim [ɾin.g-im]	‘my boxing ring’
c.	p ~ b	kap [kɑp]	‘container’	kabım [kɑ.b-ɯm]	‘my container’
	t ~ d	tat [tɑt]	‘taste’	tadım [tɑ.d-ɯm]	‘my taste’
	ç ~ c [tʃ ~ dʒ]	güç [gytʃ]	‘strength’	gücüm [gy.dʒ-ym]	‘my strength’
	k ~ g	renk [ɾenk]	‘color’	rengim [ɾen.g-im]	‘my color’

This phenomenon is not exclusive to roots. There are also some alternating and non-alternating final suffix plosives, though they are not very numerous (see the end of subsection 2.6.1 and table 63).

In addition, there are a number of plosive onsets in suffixes which alternate, as well as those which do not. The alternating plosives in initial suffix onsets always match the preceding sound in laryngeal setting. For example, the locative is one such suffix which has an alveolar stop in its initial onset: compare *evde* [ev.-de] ‘at home’ and *sokakta* [so.kɑk.-tɑ] ‘in the street’. Not all initial suffix plosives alternate. Some, such as the converbial suffix, never change in laryngeal setting: *okurken* [okuɾ-ken] ‘while reading’, cf **okurgen* *[okuɾ-gen].

The first step of the analysis given below will be to argue that aspiration is the phonologically salient feature for the laryngeal contrast in Turkish plosives. Following that, plosives will be sorted into four types: non-alternating aspirated plosives, alternating coda plosives, non-alternating plain coda plosives, and alternating suffix-initial plosives.

Representations for each of the above categories will be given using EFP directed graphs, with a new and innovative representation being proposed for non-alternating plain coda plosives.

6.1 Representing the Turkish plosive laryngeal contrast with EFP

Turkish has a two-way laryngeal contrast in its consonants. Phonetically, Turkish has been described as having a voiced stop series with prevoicing, and a voiceless stop series which is enhanced with aspiration (see Göksel & Kerslake 2006, Chapter 1 and Öğüt et al. 2006). This leaves open the question as to which feature (aspiration or voicing) is phonologically contrastive.

It seems likely that for plosives, Turkish is an aspirating language with phonetic voicing on unmarked stops. I am not the first to suggest that [voice] is not contrastive for Turkish plosives, Avery (1996) also argued that [voice] is not a phonological feature of Turkish plosives. There are several reasons which suggest this to be the case. First, there is unreliability of phonetic voicing on *g*, which has been documented in several studies of Turkish plosives (Kallestinova 2004, Öğüt et al. 2006, Kılıç 2018). Secondly, Turkish has no regressive voicing assimilation, which is a characteristic of languages with a phonological contrast involving [voice] (van Rooy & Wissing 2001). Thirdly, there is the interesting finding by Kılıç (2018) which is that native Turkish speakers seem fairly able to modulate their prevoicing on <b d g> when speaking English. This could indicate that prevoicing is not a contrastive feature, and is an optional enhancement which serves to emphasize the aspiration contrast, and can thus be more easily omitted at will. So, phonologically, the Turkish plosives are as below:

(171) *Turkish plosive phonemic inventory*

Orthography	*IPA*
Affricates:	
ç c	/$ʧ^h$ ʧ/
Stops:	
p t k	/p^h t^h k^h c^h/
b d g	/p t k c/

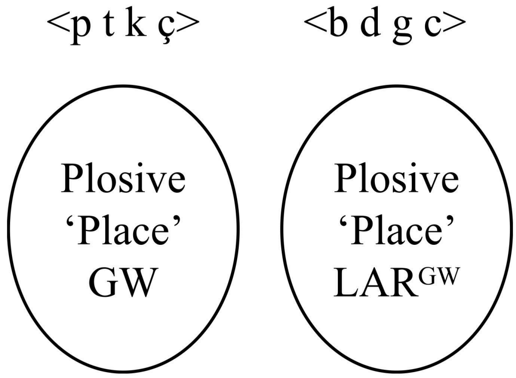
Phonetically, however, studies have found that Turkish plain plosives are often voiced, but may sometimes be pronounced plain, especially the velar stop (Kallestinova 2004, Öğüt et al. 2006, Kılıç 2018). The plosives are typically pronounced as follows:

(172) *Turkish plosive phones*

Orthography	*IPA*
Affricates:	
ç c	[$ʧ^h$ dʒ]
Stops:	
p t k	[p^h t^h k^h c^h]
b d g	[b d g ɟ]

This dissertation adopts Laryngeal Realism (see Avery & Idsardi 2001 and Iverson & Salmons 1995). As such, the ‘voiceless’ plosive series <p t k ç> is marked with a Glottal Width (GW) dimension, which will be pronounced with a [spread] gesture. This [spread] gesture opens the glottis, allowing air to pass through and facilitating aspiration. The plain plosives <b d g c> have a bare GW dimension node, Larynx^{GW} (Lar^{GW}). See below for examples of aspirated plosives (left) and plain plosives (right).

(173) *Example consonant nodes*



In these examples, the place of articulation has been left with a placeholder ('Place'), but in the actual phonological representation, this would be filled in with a place (e.g., alveolar, velar, etc). These events are given for illustrative purposes, the actual representations for stops in Turkish will end up looking slightly different, to account for the laryngeal setting alternation in the data.

The presence of phonetic voicing on plain plosives <b d g c> is represented using an enhancement rule. Enhancement rules are a part of phonetics, not present in the phonology, and are used to overmark contrasts in the surface form (see section 2.2 of the literature review for more information). Here, an enhancement rule often applies to the plain series causing voicing. This rule is given below.

(174) *Enhancement rule for plain consonants*

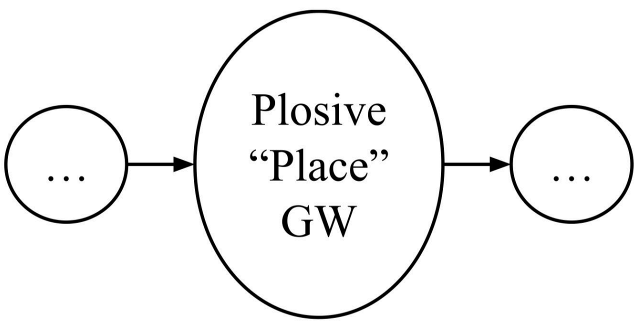
Larynx^{GW} (LAR^{GW}) → [slack] (slack vocal folds, facilitating voicing)

As mentioned in the literature review, there are two main approaches to laryngeal alternation in Turkish plosives: an underspecification approach and a coda devoicing approach. This account will combine parts of each to provide a full account of these alternations. To this end, I have separated the plosives into four groups: non-alternating aspirated plosives, alternating coda plosives, non-alternating plain coda plosives, and alternating suffix-initial plosives.

6.2 Non-alternating aspirated plosives

The first type of plosive relevant for this discussion are plosives which are non-alternating, and always aspirated (in past accounts which have analyzed Turkish as having a voiced/voiceless distinction, these would be 'always voiceless'). These are plosives like the final [t^h] in *at* [$ɑt^h$] 'horse', which remains aspirated when resyllabified into an onset, eg *atım* [a.t^h-ɯm] 'my horse'. This category also includes the always aspirated final suffix plosives, like the forms of the causative that end in a non-alternating *t* (*-t*, *-It*, *-Art*) and the non-alternating *k* in the converbial suffix *-ken* (e.g., *okurken* [ok^hurk^hen] 'while reading'). These are the most simple to account for: these plosives are underlying aspirated. In our framework, these plosives have a GW feature which marks them for aspiration inside the segment, as shown in the graph below. This is the same representation given for aspirated plosives in (173) above.

(175) *PROP graph for always aspirated plosives*



6.3 Coda plosive fortition rule / alternating root final plosives

Some grammars report that Turkish disallows ‘voiced’ plosives <b d g c> in coda position (Kornfilt 1997:491, Underhill 1976:41, Göksel & Kerslake 2006:15, see also Özçelik 2024). This account takes the stance that voice is not the phonologically contrastive feature for Turkish plosives; rather, aspiration is the phonologically contrastive feature. Thus, I argue (following Avery 1996) that this ‘devoicing’ rule is actually a coda fortition rule where plain plosives become aspirated in codas (see Iverson & Salmons 1995 for a similar argument regarding Germanic which also makes use of Laryngeal Realism). Adopting this adjusted version of the rule has some important benefits, as it is crucial for explaining two key points. First, it provides an explanation as to why plain plosives <b d g c> are rare in coda position in Turkish (see subsection 2.6.1 of the literature review). In most instances, the coda fortition rule applies, causing plain plosive codas to be rare.

Second, a coda fortition rule provides a way to explain why resyllabification into an onset is *crucial* for conditioning alternation in laryngeal setting for final plosives which do so. For example, the final [t] in *kanat* [kɑnɑt^{h}] ‘wing’ becomes unaspirated when resyllabified into an onset, e.g., *kanadım* [kɑnɑ.t-ɯm] ‘my wing’. However, the final plosive does not become plain when in coda position, even when next to a sonorant, e.g., *kanatlar* ‘wings’, cf **kanadlar*.

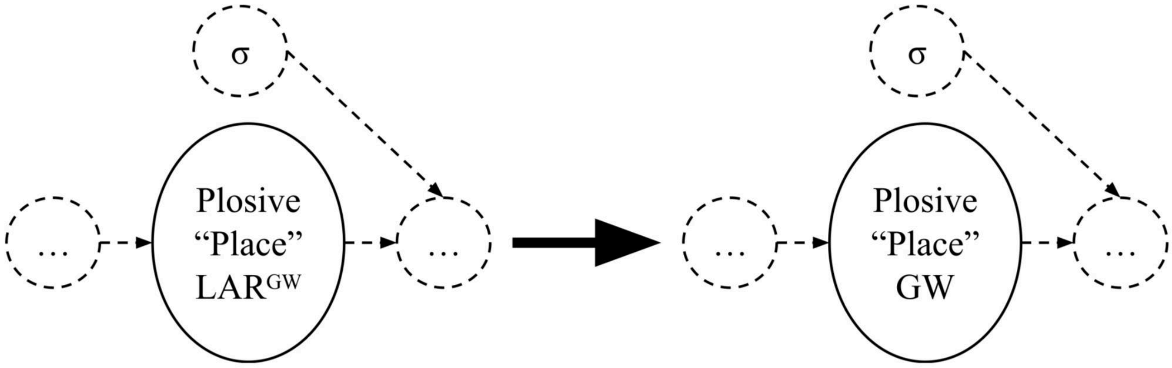
Some approaches (see Inkelas & Orgun 1995 and Petrova et al. 2006) have characterized this kind of alternating plosive as underlyingly unspecified for voicing, and the alternation arising from passive voicing through the unspecified plosive. However, adopting this view means that the passive voicing in Turkish plosives must distinguish between [+sonorant +consonant]_[vowel] and [vowel]_[+sonorant +consonant] environments. For example, the <p> in *kalp* ‘heart’ alternates in laryngeal setting when a vowel-initial suffix is added, as in *kalb-im* ‘my heart’. This means that so-called passive voicing applies in a

[+sonorant +consonant]_[vowel] environment. Recall, however, that the alternating <t> in *kanatlar* 'wings' is does not alternate in a [vowel]_[+sonorant +consonant] environment (cf., **kanadlar*). If this alternation is truly a result of phonetic passive voicing, it seems odd that the order of the sonorants (consonant or vowel being first) should be crucial for alternation.

Positing a coda fortition rule works well with this data: in the bare form, the final plosive is in coda position, so a coda fortition rule applies, yielding an aspirated plosive (e.g., *kanat* [kɑ.nɑt^h] 'wing'). However, when the final coda plosive is resyllabified into an onset, this rule fails to apply (e.g., *kanadım* [kɑ.nɑ.t-ɯm] 'my wing'). And, when a consonant initial suffix is added, even if that consonant is a sonorant like [l]- no alternation would be expected, since the final coda consonant is not resyllabified in that case (*kanatlar* 'wings', cf **kanadlar*)

Below is a representation of this coda fortition rule in an EFP graph structure. The graph on the left shows the underlying representation, with a Lar^{GW} (plain) plosive in coda position (see section 4.1 for more details on encoding syllabification in EFP graphs). The coda fortition rule checks if a Plosive event is occurring in coda position, and fills in the bare Lar^{GW} superordinate marking with a GW dimension. This changes the plosive from plain to aspirated. In other words, the coda fortition rule disallows Plosive and Lar^{GW} from occurring in the same event in coda position. Note that it is crucial for the fortition rule to check that the event is a Plosive specifically, as fricatives can appear in codas regardless of laryngeal setting, and do not display alternation.

(176) *Coda fortition rule EFP*



The above graph representation can also be represented using a more conventional phonological rule in the following way:

(177) *Plosive coda fortition rule*

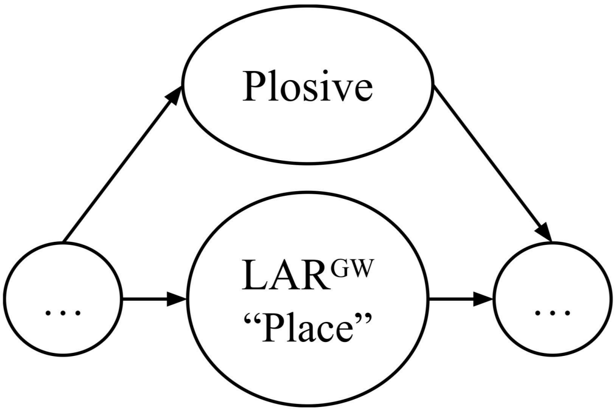
/plosive, $Lar^{GW}/_{syll}$ → [plosive GW]

6.4 Non-alternating final plain plosives

In the previous subsection, I laid out a coda fortition rule which applies to plain plosives. There remains an important issue, however: how to deal with plain coda plosives which do not receive final fortition. These can occur in root-final position (e.g., *ad* [ɑt] ‘name’) or word-medially (e.g., *abla* /apla/ ‘older sister’). Clearly the final fortition rule must not apply to these codas. However, it is not satisfactory to simply label these ‘exceptions’ and move on. It is important to take seriously how these ‘exceptions’ are encoded, phonologically. To this end, I propose that for the plain plosives in coda position, their Plosive feature is stored in a parallel autosegmental stream. Recall that the coda devoicing rule checks for Plosive and Lar^{GW} co-occurring *in the same event*. In this case, since Plosive is not contained in the event itself and therefore does not condition final fortition. When the word is pronounced, the plosive dimension will be phonetically sequenced as if it is 'in' the consonant event, leading to a plain plosive in coda position.

This proposal is fundamentally very similar to the proposal made by Avery (1996). Avery (1996) argued that these plosives are 'Sonorant Voice', because sonorants are not affected by coda fortition. This idea that these plosives are somehow phonologically not being grouped with typical plosives, and thus not subject to final fortition, is central to this analysis.

(178) *Non-alternating plain plosive in coda*



The above graph may seem odd or unnatural, as autosegmental phonology is typically used to represent non-local phonological phenomena, like harmony systems, and no such non-local phonology is present here. However, recall that this type of plosive is very uncommon Turkish- plain plosives <b d g c> are rarely found in codas. For this reason, it makes sense that the representation should be a bit 'unnatural' or difficult to maintain.

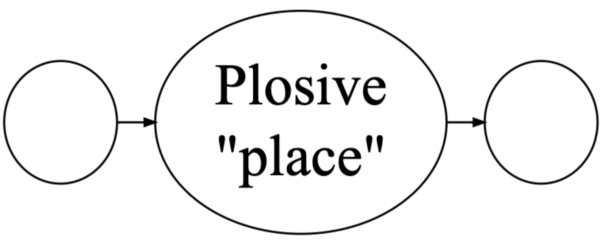
In some cases, it has been reported that Turkish speakers regularize forms with plain plosives in coda position (Lewis 1976, Özçelik 2024). In this case, the 'Plosive' manner has been de-autosegmentalized and reincorporated into the consonant event. When this occurs, the plosive is subject to final fortition.

6.5 Alternating suffix initial plosives

The final case which remains to be dealt with are suffix-initial plosives. Some suffixes have initial plosives which are plain following phonetically voiced or plain phonemes and aspirated following aspirated phonemes (e.g., sacda [sɑtʃ-tɑ] 'on the sheet metal', *ev**d**e* [ef-**t**e][24] 'at home' and *sokak**t**a* [sokʰɑkʰ-**tʰ**ɑ] 'in the street'). The coda fortition rule obviously cannot account for the laryngeal alternation in this case, because these plosives are always in onset position.

As a result, I combine the underspecification approach with the coda fortition approach given above. These alternating plosives *always* follow the same pattern based on assimilating to the laryngeal setting of the preceding sound, so I argue that these plosives are truly unspecified for voicing. When in the environment of [aspirated sound]_[Vowel], the aspirated gesture is spread onto the following sound, leading to a phone that is phonetically aspirated. When in the environment of [plain obstruent]_[Vowel] or [sonorant]_[Vowel], no aspiration is present, and the unspecified plosive remains plain, or possibly passively voiced. As shown below, the Plosive event for the alternating suffix initial plosives has no specification for laryngeal setting in the phonology.

(179) *Alternating suffix plosives*



---

[24] I do not have the space to discuss fricatives in-depth here, but I am suggesting in this transcription that fricatives in Turkish also have an aspirated/plain distinction. Another possibility is that, as Avery (1996) suggests, while plosives in Turkish are phonologically aspirated/plain, fricatives have a voiced/voiceless distinction in Turkish. Either way, this analysis would work correctly.

6.6 Conclusion

This chapter has laid out a brief account of voicing alternation in Turkish plosives using the EFP framework. My account combines both an underspecification approach and coda devoicing. So-called 'coda devoicing', however, is re-analyzed as coda fortition. In short, non-alternating aspirated plosives are underlyingly aspirated. Alternating final plosives are underlyingly plain, but aspirated in coda position due to a coda fortition rule. Non-alternating plain final plosives have a 'floating' plosive manner dimension which is stored in parallel to the actual consonant event. Since the plosive dimension is not inside the consonant event, these plosives are not subject to the coda fortition rule. Finally, alternating plosives in suffix onsets are underlyingly unspecified, receiving a laryngeal setting based on the preceding sound.

7 Conclusion

This section contains a brief summary of the contents of this work, as well as the number of areas left open for further research.

7.1 Goals achieved

The primary goal of this thesis was to find novel and innovative ways to represent ‘exceptions’ in phonology. There exist many well-known ways to account for regular patterns in speech sounds, such as phonological rules and constraints. Language data which do not fit patterns are generally much more difficult to deal with. To this end, a number of topics related to ‘exceptions’ in Turkish phonology were selected and analyzed.

Further specifics are summarized below, but the general philosophy of this dissertation has been not to ‘explain away’ surprisingly or seemingly contradictory data, but rather to show how ‘exceptions’ can inform linguists about the nature of phonological representations in the ‘regular’ case.

A secondary goal, achieved in the process of accomplishing the above, was to apply the *Events, Features, and Precedence* (EFP) framework in-depth to a particular language. The most comprehensive work which made use of EFP principles was Papillon’s (2020) dissertation. That work was ambitious, and demonstrated that EFP-type representations were useful for analyzing a number of phonological phenomena across the world’s languages. This work has built on that one by demonstrating that EFP can be applied to various phonological phenomena in a particular language, in this case Turkish, with great success.

There were four main subtopics explored in this dissertation: vowel harmony, vowel alternation in coda clusters, emphatic reduplication, and plosive laryngeal alternation. Each will be listed and summarized here.

Vowel harmony (in terms of frontness and rounding) is one of the most well-known phonological phenomena of the Turkish language. However, there are numerous exceptions to Turkish vowel harmony, both in terms of immutable or unchanging vowels in roots and affixes, and also in terms of otherwise harmonizing suffixes which do not seem to conform to frontness harmony when affixed certain roots. In order to represent these exceptions, additional autosegmental harmony event(s) are added into parallel autosegmental stream(s). The ability to add more than one TT (frontness) or LAB (rounding) event into a given autosegmental stream is already necessary for representing rounding opacity, a natural consequence of how vowel harmony operates in Turkish, even for words that conform to TVH. Therefore, the representation of 'exceptions' does not introduce any new representations into the phonology.

Next, there is the topic of a ∅ ~ high vowel alternation which occurs in some coda clusters in Turkish. Many of the alternations occur in words which would otherwise have illicit coda clusters (e.g., *adil* [ɑdil] 'justice' ~ *adli* [ɑdli] 'justice-ACC'; [dl] is not attested as a coda cluster in Turkish). This has led some researchers to analyze this as vowel epenthesis which occurs to repair illicit clusters (Lewis 1967, Clements & Sezer 1982, Kornfilt 1986, Kabak 2011, Ketrez 2012, Özçelik 2024). However, there are some problematic 'exceptional' forms which undergo alternation despite not having an illicit cluster (e.g., *nakit* [nɑkit] 'cash' ~ *nakti* [nɑkti] 'cash-ACC'; c.f. *pakt* [pɑkt] 'pact'). The account outlined in this work utilizes a 'dangling' vowel event on the underlying representations of words which undergo the alternation. This event is 'dangling' because it follows nothing. The vowel event remains dangling (and thus not pronounced) when the event following it is in an onset, and is sequenced (and thus pronounced) when the event following it is in a coda. This avoids the issue with the epenthesis account outlined earlier.

Thirdly, emphatic reduplication is a partial prefixing reduplication which adds an unpredictable linker consonant from the set <p m s r> (e.g., *pembe* [pembe] 'pink' → *pespembe* [pe-s-pembe] 'extremely pink'). In some cases, there are highly exceptional forms which take an additional added vowel or additional added vowel and consonant (e.g., *çıplak* [tʃɯplɑk] 'naked' → *çırılçıplak* [tʃɯ-ɾɯl-tʃɯplɑk] 'stark naked'). This kind of reduplication is not productive in Turkish, applying to only a limited set of words. Various accounts have been put forward regarding the selection of the linker consonant which center on dissimilation or identity avoidance with consonants of the base (Taneri 1990, Wedel 1999, Kelepir 2001, Özçelik 2012). While these can account for general trends in the choice of linker, there are always 'exceptions' to these trends for which the linker cannot be properly predicted. In addition, novel and nonce word studies have shown that speakers are not consistent in their choice of linker on new forms (Wedel 1999, Demir 2017, Kılıç & Bozşahin 2013, Köylü 2021).

Due to these facts, this work contends that the choice of linker is lexicalized and stored as a dangling event on the underlying representation of the word. This is a different kind of dangling event than the one which stores the dangling vowel for ∅ ~ high vowel alternation. In this case, the dangling event precedes nothing. When TER is applied to a word, the dangling link event is sequenced before the first event of the word. This structure can account for the linker which appears on every word which receives TER, and also serves to mark which words can undergo it (recall that TER only applies to a limited set of words).

The final topic covered in this dissertation is plosive laryngeal alternation. In Turkish, final plosives in roots and suffixes can alternate in terms of laryngeal setting when resyllabified from a coda into an onset. In addition, some suffix initial plosives alternate in laryngeal setting depending on the laryngeal setting of the preceding sound. Some phonologists have argued that this reflects an underlying three-way voicing contrast where always voiced plosives are underlyingly voiced, always voiceless plosives are underlyingly voiceless, and alternating plosives are underlyingly unspecified (Inkelas & Orgun 1995, Kellestinova 2004, Petrova et al. 2006). Another approach is to posit a final or coda devoicing rule and analyze alternating final plosives as being devoiced (Underhill 1976, Kornfilt 1997, Göksel & Kerslake 2006, Özçelik 2024). In this account, always voiced final plosives are 'exceptions' to coda devoicing.

Based on phonological and phonetic evidence, I argue that [voice] is not a contrastive feature in Turkish; rather, it is aspiration or [spread] which is contrastive (Avery 1996 also argues that this is the case). In the account provided here, always [spread] plosives are underlyingly GW. Alternating final plosives are subject to a coda fortition rule which applies when 'plosive' and 'LAR^{GW}' co-occur in a coda event. Final plosives which surface as LAR^{GW} do so because their manner 'plosive' is stored parallel to the event, in an autosegmental stream. Thus 'plosive' and 'LAR^{GW}' do not occur in the same event, and coda fortition does not apply. This idea is fundamentally very similar to the analysis of this issue by Avery (1996), which made use of Sonorant Voice for these plosives. Finally, alternating suffix-initial plosives are truly unspecified for laryngeal setting, always matching laryngeal setting of the preceding sound. This idea combines the two past approaches of coda devoicing and underspecification. It also accounts for the 'exceptions' to coda devoicing without the need for any exception diacritic.

7.2 Areas for future research

There are many issues related to the topics discussed in this work which would benefit immensely from further research.

There are several issues yet to be addressed related to vowel harmony. One is how, exactly, the consonants are affected by the vowel harmony autosegmental streams. The consonants /k g l/ have front and back allophones, and which appear in the surface form is related to the frontness of the surrounding vowels (Clements & Sezer 1982, Göksel & Kerslake 2006). This work did not delve very much into how the front and back allophones are conditioned based on the autosegmental harmony information. Aside from this, when looking at exceptions to TVH, it is not clear why in modern İstanbul Turkish front vowel roots which take exceptional back vowel suffixes are not attested, but there remain in use many back vowel words which take front vowel suffixes (e.g., *saati* [sɑːt-i] 'clock-ACC', *rolü* [rol-y] 'role-ACC', etc).

Furthermore, this work (and many other accounts of TVH) focused specifically on the İstanbul dialect of Turkish. Recent promising work has been done by Demir (2023) documenting vowel harmony in Laz Turkish, which is a dialect spoken by the minority Laz ethnic group in the Black Sea (Karadeniz) region of Turkey. More work is needed to incorporate the vowel harmony patterns in Laz Turkish and other minority dialects of Turkish into phonological accounts of TVH.

A sketch of syllable structure representation in the EFP framework was given as a foundation for the analysis of vowel alternation in coda clusters (since the analysis required reference to what is a coda and an onset in an EFP graph). The representation of syllable structure is preliminary, and would likely benefit from further scrutiny and improvement. For example, in this work, syllable events preceded onsets. However, there are many other ways in which syllable structure could be encoded. For example, perhaps it is the nuclei (vowels) which are preceded by syllable events. The syllable structure given in this work is meant to be a starting point from which future work on syllable structure in EFP can grow.

In terms of emphatic reduplication (TER), there are many interesting areas for further study. As discussed in Chapter 5, the selection of the linker consonant on attested forms is not a result of a phonological process, but rather, the choice is specified in the lexicon. However, there is still the matter of 'fuzzy' patterns in the selection of the linker on novel forms. I have pointed out some non-phonological strategies (mental foraging, probability matching) which could be appropriate in accounting for patterns, along with variation (see subsection 5.4.1). To my knowledge, there is only one piece of literature on how the linker may be chosen on novel forms via related non-phonological means (Kılıç & Bozşahin 2013). More research is needed to investigate how non-phonological strategies are employed in the selection of linker for novel forms.

In addition, socio-linguistic variation in linkers could be a promising future research direction. There have been some suggestions that dialect may affect the selection of linker (Taneri 1990), as well as one study on Azerbaijani Turkish (Abbasi & Moradkhani 2012), which found that linker selection was different for that dialect. However, another study (Demir 2017) found that dialect did not have an effect on linker choice. So dialectical variation in linker choice would be an interesting area for future research. Finally, an examination of sociolinguistic attitudes towards the selection of linker would be insightful. In private communication with native Turkish speakers, they have stated that they associate the selection of certain linker consonants and vowels with sociolinguistic factors, such as age. It would be interesting to see whether this is indicative of wider overall sociolinguistic attitudes around this form of reduplication.

The chapter on laryngeal alternation in plosives notably excluded an analysis of velar drop (the *k* ~ *ğ* alternation often observed a morpheme boundary). Velar drop seems linked to laryngeal alternation in that rather than alternating with g in most environments, *k* alternates with *ğ* instead. The two phonemes *ğ* and *g* were at least historically linked, being in complementary distribution in native Turkish words. However, in the current day, the two are no longer in complementary distribution due to the introduction of loanwords (Ünal-Logacev 2019). This related topic would benefit from a phonological analysis which either incorporates it into the broader pattern of laryngeal alternation in Turkish plosives, or separates it out and accounts for it as a separate phonological phenomenon.

Finally, this dissertation did not propose any analysis of clitics or compound words. Further work is needed to incorporate these into the EFP account of Turkish phonology which is given here.

There is much work left to be done. This dissertation has served as the first EFP-type in-depth phonological analysis of a particular language. More projects like this are needed to better understand the full extent of the EFP structures necessary to analyze phonology across different languages. As discussed above, there is certainly more work to be done on Turkish in particular. I hope that this dissertation can serve as a foundation for further phonological research on Turkish and beyond.

References

Abbasi, Nasser, & Elahe **Moradkhani**. (**2012**). Intensified reduplicated adjectives in Azerbaijani Turkish. *Social and Behavioral Sciences*, 70, 1600–1608.

Anselme, Patrick & Onur **Güntürkün**. (**2019**). How foraging works: Uncertainty magnifies food-seeking motivation. *Behavioral and Brain Sciences* 42(e35), 1–59.

Avery, Peter. (**1996**). *The Representation of Voicing Contrasts* (doctoral dissertation). University of Toronto.

Avery, Peter & William **Idsardi**. (**2001**) Laryngeal dimensions, completion and enhancement. In Hall, T. Alan (ed.), *Distinctive feature theory*. 41–70. Berlin: Mouton de Gruyter.

Bacanlı, Eyüp, Darin Flynn, & Amanda Pounder, (**2020**). Vowel Harmony and Other Morphological Processes in Turkish, *Proceedings of the Workshop on Turkic and Languages in Contact with Turkic* 5, 26–40. doi: https://doi.org/10.3765/ptu.v5i1.4782

Baier, Rebecca, William J. Idsardi, and Jeffrey Lidz. (**2007**). Two-month-olds are sensitive to lip rounding in dynamic and static speech events. In *Proceedings of the International Conference on Auditory-Visual Speech Processing*, L6–2. https://www.isca-speech.org/archive/avsp_2007/baier07_avsp.html

Bird, Steven & Mark **Ellison**. (**1994**). One-Level Phonology: Autosegmental Representations and Rules as Finite Automata. *Computational Linguistics* 20(1), 55–90.

Bellik, Jennifer. (**2018**). An acoustic study of vowel intrusion in Turkish onset clusters, *Laboratory Phonology* 9(1):16, 1–23. doi: https://doi.org/10.5334/labphon.112

Canalis, Stefano & Furkan **Dikmen**. (**2020**). *Proceedings of the Workshop on Turkic and Languages in Contact with Turkic* 5, 41–55. https://doi.org/10.3765/ptu.v5i1.4781

Cho, Taechong, D.H. Whalen, & Gerard Docherty. (**2019**). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics* 72, 52–65.

Clements, George N., & Engin **Sezer**. (**1982**). Vowel and consonant disharmony in Turkish. *The structure of phonological representations* 2, 213–255.

Demir, Neşe. (**2017**). *Turkish Reduplicative Adjectives and Adverbs* (master's thesis). Syracuse University.

Demir, Neşe. (**2018**). Turkish reduplicative adjectives and adverbs. *Proceedings of the Linguistic Society of America*, 3(1), 19–1.

Demir, Neşe. (**2023**). *Laz Turkish: A case study of partially productive vowel harmony and sociolinguistic attitudes* (doctoral dissertation). University of California, San Diego.

Demircan, Ömer. (**1987**). Emphatic reduplications in Turkish. In Boeschoten, H. E. & Verhoeven, L. T. (eds.), *Studies in modern Turkish: Proceedings of the third conference on Turkish Linguistics*, 24–41. Tilburg: Tilburg University Press.

Dil Derneği. (n.d.). rol. In *Türkçe Sözlük*. Retrieved March 3, 2026, from https://www.dildernegi.org.tr/TR,274/turkce-sozluk-ara-bul.html.

Dobrovolsky, Michael. (**1987**). Why CVC in Turkish reduplication? *The Nordic languages and modern linguistics 6: Proceedings of the sixth International Conference of the Nordic and General Linguistics, Helsinki*. Helsinki: Helsinki University Press.

Dresher, B. Elan. (**2008**). *The contrastive hierarchy in phonology.* Contrast in Phonology: Theory, Perception, Acquisition. doi: https://doi.org/10.1017/cbo9780511642005

Erdal, Marcel. (**2010**). Inalienability and syncopation in Turkish. In Hendrik Boeschoten & Julian Rentzsch (eds.), *Turcology in Mainz - Turkologie in Mainz.* 147–153. Wiesbaden: Harrassowitz.

Feizollahi, Zhaleh. (**2009**). *Two Case Studies in the Phonetics~Phonology Interface: Evidence from Turkish Voicing and Norwegian Coalescence*. Dissertation, Georgetown University.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen, & Kevin Russell. (**2004**). Contrastive Focus Reduplication in English (The Salad-Salad Paper). *Natural Language & Linguistic Theory* 22, 307–357

Goldsmith, John. (**1976**). *Autosegmental phonology.* (Doctoral dissertation, MIT Press).

Göksel, Aslı & Celia **Kerslake**. (**2006**). *Turkish: A Comprehensive Grammar*. London: Routledge.

Hayes, Bruce. (**2009**). *Introductory phonology*. Vol. 7. John Wiley & Sons.

Hegalsen, Pétur & Catherine **Ringen**. (**2008**). Voicing and aspiration in Swedish stops. *Journal of phonetics*, 36(4), 607-628. doi: 10.1016/j.wocn.2008.02.003

Idsardi, William. (**2022**). Underspecification in time. *Canadian Journal of Linguistics*, 67(4), 670–682.

Inkelas, Sharon. & Cemil Orhan **Orgun** (**1995**). Level ordering and economy in the lexical phonology of Turkish. *Language* 71(4), 763–793.

Iverson, Gregory K. & Joseph C. **Salmons**. (**1995**). Aspiration and Laryngeal Representation in Germanic. *Phonology* 12(5), 369–396. http://www.jstor.org/stable/4420084?origin=JSTOR-pdf

Kabak, Barış. (**2011**). Turkish Vowel Harmony. In: Marc van Oostendorp, Colin Ewen, Beth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*. Wiley-Blackwell.

Kabak, Barış, Eva Kasselkus, Kazumi Maniwa & Silke Weber. (**2008**). *Vowel harmony has direction and context: A corpus study*. Paper presented at the 16th Manchester Phonology Meeting.

Kabak, Barış & S. **Weber** (**2013**). Markedness, context, and directionality in vowel harmony: A corpus study on vowel co-occurence patterns. *Dilbilim Araştırmaları* [Journal of Linguistic Research] 2013, 53–85.

Kaufman, Brianna Danielle. (**2014**). *Learning an unproductive process: Turkish emphatic reduplication* (master's thesis). UC Santa Cruz. https://escholarship.org/uc/item/8mp2d399

Kallestinova, Elena. (**2004**). Voice and aspiration of stops in Turkish, *Folia Linguistica* 38(1–2), 117–43.

Kaun, Abigail R. (**1999**). Epenthesis-Driven Harmony in Turkish. *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society: Special Session on Caucasian, Dravidian, and Turkic Linguistics* (2000), 95–106.

Kelepir, Meltem. (**2001**). To be or not to be faithful. In Göksel A., & Kerslake, C. (eds.), *Studies on Turkish and Turkic Languages: Proceedings of the 9th International Conference on Turkish Linguistics*, 12-14 August 1998, Oxford, 11–18. Wiesbaden: Harrassowitz.

Ketrez, F. Nihan. (**2012**). *A student grammar of Turkish*. Cambridge University Press.

Kılıç, Mehmet. (**2018**). Voice onset time as an indicator of interlanguage: Evidence from Turkish-accented English. *Journal of Language and Linguistic Studies*, 14(4), 272–284.

Kılıç, Özkan & Cem **Bozşahin**. (**2013**). Selection of Linker Type in Emphatic Reduplication:

Speaker's Intuition meets Corpus Statistics. *Proceedings of the Annual Meeting of the Cognitive Science Society* 35(35), 2722–2727. https://escholarship.org/uc/item/24r2b2cd

Kirchner, Robert. **(1993)**. *Turkish vowel disharmony in Optimality Theory*. Paper presented at the Rutgers Optimality Workshop 1, Rutgers University.

Kornfilt, Jaklin. (**1997**). *Turkish* (1st ed.). Routledge. https://doi.org/10.4324/9781315823652

Kornfilt, Jaklin. (**1985**). Stem-Penultimate Empty Cs, Compensatory Lengthening, and Vowel Epenthesis in Turkish. In Leo Wetzels & Engin Sezer (eds.), *Studies in compensatory lengthening* (pp 79–96). Foris Publications.

Köylü, Yılmaz. (**2021**). Abstract knowledge of emphatic reduplication in Turkish. In *Proceedings of the Workshop on Turkic and Languages in Contact with Turkic*, 5(1), 86–96.

Lewis, Geoffrey. (**1967**). *Turkish Grammar*. Oxford: Clarendon Press.

Li, Xin & Eric **Raimy**. (**2017**). Laryngeal realism for Mandarin tones. *Beyond VOT - Searching for Realism in Laryngeal Phonology*. Poznan Linguistics Meeting, Poznan, PL.

Lightner, Theodore M. (**1965**). On the Description of Vowel and Consonant Harmony. *Word* 21(2), 244–250, doi: 10.1080/00437956.1965.11435427

Lisker, Leigh & Arthur **Abramson**. (**1967**). Some Effects of Context on Voice Onset Time in English Stops. *Language and Speech* 10(1), 1–28.

Lundin, Nancy B., Joshua W. Brown, Brendan T. Johns, Michael N. Jones, John R. Purcell, William P. Hetrick, Brian F. O'Donnell, & Peter M. Todd. (**2023**). Neural Evidence of switch processes during semantic and phonetics foraging in human memory. *Proceedings of the National Academy of Sciences* (PNAS) 120(42), 1–12.

Neary, Terrance M. & Bernard L. **Rochet** (**1994**). Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association* 24(1), 1–18.

Nevins, Andrew. (**2010**). *Locality in Vowel Harmony*. The MIT Press.

Nicolae, Andreea C & Andrew **Nevins**. (**2016**). Fricative patterning in aspirating versus true voice languages. *Journal of Linguistics* 52, 151–174. doi: 10.1017/S0022226715000067

Meral, Semra Baturay. (**2023**). The Medial and Final (CCC) Consonant Clusters in the Loanwords of Turkish: Is It an Illusion?. *IBAD Journal of Social Sciences* 15, 352–373. doi: https://doi.org/10.21733/ibad.1367613

Odden, David. (**2005**). *Introducing Phonology*. Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511808869

Öğüt, Fatih, Mehmet Akif Kiliç, Erkan Zeki Engin, & Raşit Midilli. (**2006**). Voice onset times for Turkish stop consonants. *Speech Communication* 48, 1094–1099.

Özçelik, **Sadettin**. (2012). m, p, r, s Ünsüzleriyle Yapılan Pekiştirme ve Kuralları Üzerine. *Belletin* 2, 29–42.

Özçelik, Öner. (**2024**). *The Phonology of Turkish*. Oxford University Press.

Papillon, Maxime. (**2020**). *Precedence and the lack thereof: Precedence-relation-oriented phonology* (doctoral dissertation). University of Maryland, College Park.

Petrova, Olga, Rosemary Plapp, Catherine Ringen, & Szilárd Szentgyörgyi. (**2006**). Voice and aspiration: Evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review* 23(1), 1–35. https://doi.org/10.1515/TLR.2006.001

Pöchtrager, Markus Alexender. (**2010**). DOES TURKISH DISS HARMONY? *Acta Linguistica Hungarica*, 57(4), 458–473. http://www.jstor.org/stable/26190972

Purnell, Thomas C., Eric **Raimy**, & Joseph **Salmons**. (**2019**). Old English vowels: Diachrony, privativity, and phonological representations. *Language* 95(4), e447-e473. doi:10.1353/lan.2019.0083.

Raimy, Eric. (**2000**). *The phonology and morphology of reduplication*, volume 52. Walter de Gruyter.

Raimy, Eric. (**2021**). Privativity and ternary phonological behavior. In S. Bendjaballah, A. Tifrit & L. Voeltzel (Ed.), *Perspectives on Element Theory* (pp. 65–110). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110691948-004

Idsardi, William J. and Eric **Raimy**. (**2026**). Linearization: There Are No Strings. In Grohmann Kleanthes and Evelina Leivada (eds.), *The Cambridge Handbook of the Minimalist Program*. pp. 739–777. Cambridge University Press.

Ringen, Catherine & Kari **Suomi**. (**2012**). The voicing contrast in Fenno-Swedish stops. *Journal of Phonetics*, 40(3), 419–429. doi: 10.1016/j.wocn.2012.02.010

Stachowski, Kamil. (**2014**). *Standard Turkic C-Type Reduplications*. Jagiellonian University Press.

Sofu, Hatice. (**2005**). Acquisition of reduplication in Turkish. In Hurch, Bernhard [with editorial assistance of Veronica Mattes] (ed.). *Studies on Reduplication* Vol 28 (pp. 493–510). Mouton de Gruyter.

Sofu, Hatice, & Aslı **Altan**. (**2008**). Partial reduplication: Revisited. In S. Ay, Ö. Aydin, İ. Ergenç, S. Gökmen, S. İşsever, & D. Peçenek (eds.), *Essays on Turkish Linguistics: Proceedings of the 14th International Conference on Turkish Linguistics* (pp. 63–73). Wiesbaden, Germany: Harrassowitz Verlag.

Swift, Lloyd. (**1963**). *A Reference Grammar of Modern Turkish*. Indiana University Publications.

Taneri, Mübeccel. (**1990**). A type of reduplication in Turkish. *Kansas working papers in Linguistics*, 15, 96–132.

Tang, Kevin & Faruk **Akkuş**. (**2023**). Identity Avoidance in Turkish Partial Reduplication: Feature Specificity and Locality. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 14(1), 1–52. doi: https://doi.org/10.16995/labphon.6459

Turgay, Tacettin & Halil **İskender**. (**2021**). *Turkic Languages* 25, 107-129. DOI: 10.13173/TL.25.1.107

Türk Dil Kurumu Sözlükleri. (n.d). rol. In *Türk Dil Kurumu Sözlükleri*. Retrieved March 3, 2026, from https://sozluk.gov.tr/.

Türk Dil Kurumu Sözlükleri. (n.d). kanat. In *Türk Dil Kurumu Sözlükleri*. Retrieved February 15, 2026, from https://sozluk.gov.tr/.

Uffman, Christian. (**2011**). The Organization of Features. In van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume, & Keren Rice (eds.). *The Blackwell Companion to Phonology* (pp. 1–26). John Wiley & Sons, Ltd.

Underhill, Robert. (**1976**). *Turkish Grammar*. Cambridge: The MIT Press.

Uzun, İpek Pınar. (**2021**). Phonetic Representation of Epenthetic Glides in Turkish. In Söylemez, Ayşe Selmin & Alper Kumcu (eds.). *Synergy II: Linguistics, Contemporary Studies on Turkish Linguistics* (pp. 33–58). Peter Lang Publishing. DOI:10.3726/b18432

Ünal-Logacev, Özlem, Susanne Fuchs, & Leonardo Lancia. (**2018**). A multimodal approach to the voicing contrast in Turkish: Evidence from simultaneous measures of acoustics, intraoral pressure and tongue palatal contacts. *Journal of Phonetics* 71, 385–409. https://doi.org/10.1016/j.wocn.2018.10.002

Ünal-Logacev, Özlem, Marzena Żygis, & Susanne Fuchs. (**2019**). Phonetics and phonology of soft 'g' in Turkish. *Journal of the International Phonetic Association* 49(2), 183–206. doi:10.1017/S0025100317000317

van der Hulst, Harry. & Jeroen **van de Weijer**. (**1995**). Vowel harmony. *Handbook of phonological theory*, 495–534.

van Rooy, Bertus. & Daan **Wissing**. (2001). Distinctive [voice] implies regressive voicing assimilation. In T. Hall (Ed.), *Distinctive Feature Theory* (pp. 295–334). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110886672.295

van Schaaik, Gerjan. (**2020**). *The Oxford Turkish grammar*. Oxford University Press.

Walker, Abby. (**2020**). Voiced stops in the command performance of Southern US English. *The Journal of the Acoustical Society of America* 147(1), 606–615. DOI: 10.1121/10.0000552

Wedel, Andrew. (**1999**). *Turkish emphatic reduplication* (master's thesis). UC Santa Cruz.

Wilson, Stephen M. (**2003**). A phonetic study of voiced, voiceless and alternating stops in Turkish. *The Newsletter of the Center for Research in Language* 15(1), 1-13.

Wonnacott, Elizabeth & Elissa L **Newport**. (**2005**). Novelty and Regularization: The Effect of Novel Instances on Rule Formation. *BUCLD: Proceedings of the 29th Annual Boston University Conference on Language Development* ISBN 978-1-57473-054-8

Yılmaz, Kaan. (**2020**). Türkçede Kelime Başı Hece İkileşme ve Türeme Olayı. *Türkbilig* 40, 115–132.

Yip, Moira. (**1998**). Identity avoidance in phonology and morphology. In Steven G. Lapointe, Diane K. Brentari and Patrick M. Farrell (eds.), *Morphology and its relation to phonology and syntax*. 216–246. Stanford, CA: CSLI Publications.

Appendix 1

Word meanings from table (30):

anne	/ɑnne/	mother
armut	/ɑɾmut/	pear
asansör	/ɑsansœɾ/	elevator
böcek	/bœdʒek/	bug
cesur	/dʒesuɾ/	brave
depozit	/depozit/	deposit
dublör	/dublœɾ/	stand in; stunt double
duygu	/dujgu/	emotion
elma	/elmɑ/	apple
fritöz	/fɾitœz/	fryer
gusül	/gusyl/	full abulition
hani	/hɑni/	you know; actually
hazırun	/hazɯɾun/	participant
hırbo	/hɯɾbo/	jerk
horoz	/hoɾoz/	rooster
ışık	/ɯʃɯk]/	light
ıtriyat	/ɯtɾiyɑt/	perfumery
iki	/iki/	two
inek	/inek/	cow
ipucu	/ipudʒu/	clue
kedi	/kedi/	cat
kırlent	/kɯɾlent/	pillow
kilo	/kilo/	kilogram
kitap	/kitɑp/	book
kötü	/kœty/	bad
külot	/kylot/	panties
kubbe	/kub:e/	dome
kupa	/kupɑ/	cup
kuzin	/kuzin/	cousin
menkıbe	/menkɯbe/	legend
menü	/meny/	menu
modül	/modyl/	module
muzır	/muzɯɾ/	harmful
mücadele	/mydʒɑdele/	struggle
mücver	/mydʒveɾ/	zucchini pancake
müdür	/mydyɾ/	director
nüfus	/nyfus/	population
nöron	/nœɾon/	neuron
oda	/odɑ/	room
okul	/okul/	school

para	/paɾa/	money
poşet	/poʃet/	grocery bag
printer	/pɾintɯɾ/	printer
profesör	/profesœɾ/	professor
rodstır	/ɾodstɯɾ/	roadster
rölik	/ɾœlik/	relic
rölöve	/ɾœlœve/	(architectural) survey
rötar	/ɾœtaɾ/	delay
rötuş	/ɾœtuʃ/	touch up (as in a photo)
sarı	/saɾɯ/	yellow
sıra	/sɯɾa/	place; order
soru	/soɾu/	question
şoför	/ʃofœɾ/	chauffeur
tablo	/tablo/	table
tesadüf	/tesadyf/	coincidence
tümör	/tymœɾ/	tumor
ümit	/ymit/	hope
virüs	/viɾys/	virus
yeter	/jeteɾ/	enough

Appendix 2

See below for a table of collated TER data for consonant-initial forms. Most sources do not provide translations for the forms they provide. Where possible, I have provided a translation. Note that G & K (2006) is a shortening of Göksel & Kerslake (2006). This change was made to improve the formatting of the table. The data in this table is also hosted online and can be found at this link:

https://osf.io/f4yec/overview?view_only=8a08adc3ab7d405ba7e784a718b6705e

Bare Form	*English*	*Linker*	*Source(s)*	*Taneri (1990)*
bank		s	Yılmaz (2020)	
başka	different	m	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	
bayağı	vulgar	s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
bayat	stale	s	Wedel (1999), Stachowski (2014)	
bedava	free (of charge)	s	Wedel (1999), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
bej	beige	m	Wedel (1999)	
belli	certain	s	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Demir (2017), Yılmaz (2020)	8
beraber	together	s	Wedel (1999), Stachowski (2014), Yılmaz (2020)	
berbat	terrible	s	Köylü (2021)	
berk	solid	m	Yılmaz (2020)	
berrak	serene	s	Wedel (1999), Özçelik (2012),	

			Stachowski (2014), Yılmaz (2020)	
beter	worse	m		1
		s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	19
beyaz	white	m	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Demir (2017), van Schaaik (2020), Köylü (2021)	24
bok	despised	m	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	25
bol	abundant	m	Yılmaz (2020)	6
		s	Özçelik (2012), Yılmaz (2020)	12
boş	empty	m	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Demir (2017), van Schaaik (2020), Yılmaz (2020)	29
boz	gray, dun	m	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
budala	dumb	s	Yılmaz (2020)	
bulanık	blurry	m	Stachowski (2014), Yılmaz (2020)	
		s	Özçelik (2012), Yılmaz (2020), Stachowski (2014)	
buruş	wrinkle	m	Stachowski (2014), Yılmaz (2020)	
buruşuk	wrinkled	m	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Yılmaz (2020)	
buz	ice	m	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
bütün	complete	m		1
		s	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012),	

			Stachowski (2014), Yılmaz (2020)	25
büyük	big	s	Özçelik (2012), Yılmaz (2020)	8
canlı	lively	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Ketrez (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	
cavlak	hairless	p	Özçelik (2012), Yılmaz (2020)	
cavlak	hairless	s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
cıbıl	naked	p	Özçelik (2012), Yılmaz (2020)	
		s	Wedel (1999), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
cıbıldak	naked	p	Özçelik (2012), Yılmaz (2020)	
		s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Yılmaz (2020)	
cıbır		s	Yılmaz (2020)	
cılız	skinny	p	Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
cılk	rotten	m	Wedel (1999), Özçelik (2012), Yılmaz (2020), Stachowski (2014)	
		s	Özçelik (2012), Yılmaz (2020)	
cıvık	juicy	p	Özçelik (2012), Yılmaz (2020)	
		r	Yılmaz (2020)	
		s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
ciddi	serious	p	Köylü (2021)	
çabuk	fast	m		2
		p	Wedel (1999), Özçelik (2012), Yılmaz (2020)	
		r	Swift (1963), Wedel (1999),	

			Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Demir (2017), van Schaaik (2020), Yılmaz (2020)	29
		s	Özçelik (2012), Yılmaz (2020)	
çevre	around	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	8
çevre	around	s	Yılmaz (2020)	
çıplak	naked	p		1
		r	Wedel (1999), G & K (2006), Özçelik (2012), Ketrez (2012), Stachowski (2014), Yılmaz (2020)	25
		s	Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
çıtır	crispy	m	Yılmaz (2020)	
		s	Yılmaz (2020)	
çiğ	raw	m	Wedel (1999), Stachowski (2014), Yılmaz (2020)	
		p	Swift (1963), Özçelik (2012), Yılmaz (2020)	
çirkin	ugly	p	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
çitlem		r	Yılmaz (2020)	
çizik	scratch	p	Özçelik (2012), Yılmaz (2020)	
çukur	hollow	m	Yılmaz (2020)	
çürük	rotten	m	Yılmaz (2020), Swift (1963), Stachowski (2014)	2
		p	Özçelik (2012), Yılmaz (2020)	17
		s		2
dağınık	messy	p	Özçelik (2012), Yılmaz (2020), Stachowski (2014)	
		s	Özçelik (2012)	

dar	narrow	m	Yılmaz (2020)	
		p	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	29
		s	Özçelik (2012), Yılmaz (2020)	
daracık	skintight	p	Wedel (1999), G & K (2006), Özçelik (2012), Yılmaz (2020)	
		s	Özçelik (2012), Yılmaz (2020)	
dayı		s	Yılmaz (2020)	
dazlak	bald	m	Wedel (1999), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		p	Özçelik (2012), Yılmaz (2020)	
değirmi	circular	s	Wedel (1999), Stachowski (2014), Yılmaz (2020)	
derin	deep	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Demir (2017), Yılmaz (2020)	25
		s		1
des		m	Yılmaz (2020)	
deşik		r	Yılmaz (2020)	
dıbıl		s	Yılmaz (2020)	
dıvrak		r	Yılmaz (2020)	
		s	Yılmaz (2020)	
dızlak	bald	m	Yılmaz (2020)	
		p	Özçelik (2012), Yılmaz (2020)	
dik	steep	m	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		p	Özçelik (2012), Yılmaz (2020)	
dinç	vigorous	p	Wedel (1999), Kelepir (2001),	

			Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
diniz		p	Yılmaz (2020)	
diri	lively	m	Wedel (1999), Kelepir (2001), G & K (2006), Stachowski (2014), Yılmaz (2020)	27
		p	Yılmaz (2020)	
doğru	true	p	Özçelik (2012), Yılmaz (2020)	6
		s	Wedel (1999), G & K (2006), Ketrez (2012), Özçelik (2012), Yılmaz (2020), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	26
dolayı	due to	p	Yılmaz (2020)	
		s	Yılmaz (2020)	
dolu	full	p	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	30
		s	Stachowski (2014)	
donuk	dull	p	Özçelik (2012), Yılmaz (2020)	
durgun	stagnant	p	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020), Köylü (2021)	
duru	clear	m	Yılmaz (2020)	
		p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
düdük		m	Yılmaz (2020)	
dürü	roll(ed up)	p	Wedel (1999), Yılmaz (2020), Stachowski (2014)	
düz	flat	m	Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		p	Wedel (1999), Ketrez (2012),	

			Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
düzen		p	Yılmaz (2020)	
düzenli	tidy	p	Özçelik (2012), Yılmaz (2020)	
düzgün	straight	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
gece	nightly	p	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
genç	young	p	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	24
geniş	wide	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Ketrez (2012), Stachowski (2014), Yılmaz (2020)	21
gergin	stressed	p	Wedel (1999), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	8
gevrek	crunchy; brittle	p	Özçelik (2012), Yılmaz (2020)	
		s		5
gök	sky [blue]	m	Özçelik (2012), Yılmaz (2020)	
		p	Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
götürü	bunched	s	Stachowski (2014), Yılmaz (2020)	
gri	gray	p	Köylü (2021)	10
		s		1
güdelek		s	Yılmaz (2020)	
güdük	stubby	s	Stachowski (2014), Yılmaz (2020)	
gündüz	by day	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
gür	bushy	m		2

		p	Wedel (1999), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	17
gürültülü	noisy	p	Köylü (2021)	
güzel	beautiful	m		1
		p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	15
		s		4
hazır	prepared	p	Yılmaz (2020)	
heyecanlı	exciting	p	Köylü (2021)	2
hışır	stupid	r	Yılmaz (2020)	
hızlı	fast	p	Wedel (1999), Yılmaz (2020), Köylü (2021)	
kahve	coffee	p	Wedel (1999)	
kahverengi	brown	p	Köylü (2021)	2
kalın	thick	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	25
kalın	thick	s		4
kapalı	closed	p	Wedel (1999)	
kar		m	Wedel (1999)	
kara	dark	p	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	27
		s	Yılmaz (2020)	3
karanlık	shadowy	p	Wedel (1999), Kelepir (2001), G & K (20060, Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	29
		s	Yılmaz (2020)	1
katı	stiff	p	Özçelik (2012), Stachowski (2014), Köylü (2021), Yılmaz (2020)	7

		s	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020), Köylü (2021)	22
kel	bald	p	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
kırmızı	red	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020), Köylü (2021)	23
		s	Köylü (2021)	1
kısa	short	p	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
kısacık	short	p	Özçelik (2012), Yılmaz (2020)	
kısık	hoarse	p	Özçelik (2012), Yılmaz (2020)	
kıvrak	agile	p	Özçelik (2012), Yılmaz (2020)	
		s	Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
kızıl	red	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020), Köylü (2021)	22
		s	Wedel (1999)	1
kirli	dirty	p	Wedel (1999), G & K (2006), Stachowski (2014), Özçelik (2012), Yılmaz (2020), Köylü (2021)	
koca	huge	p	Özçelik (2012), Yılmaz (2020)	
		s	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	25
kocaman	huge	p	Özçelik (2012), Yılmaz (2020)	2

		s	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	25
kolay	easy	p	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		s	Özçelik (2012), Yılmaz (2020)	
komik	funny	s	Wedel (1999)	
koyu	dark; strong	p	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	24
		s	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	7
kör	blind	m	Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		p	Özçelik (2012), Yılmaz (2020)	
kötü	bad	p	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	8
		s	Wedel (1999), Özçelik (2012), Yılmaz (2020)	9
kötürüm	crippled	p	Özçelik (2012), Yılmaz (2020)	
		s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski 2014, Yılmaz 2020	
kuru	dry	p	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	22
küçük	small	m		5
		p	Özçelik (2012), Yılmaz (2020)	10
		s	Özçelik (2012), Yılmaz (2020)	6
kütük	log (redup. form means 'completely	p	Özçelik (2012), Yılmaz (2020)	

	drunk')	s	Wedel (1999), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	
lacivert	navy blue	p	Wedel (1999), Stachowski (2014), Yılmaz (2020), Köylü (2021)	
		s	Köylü (2021)	
mavi	blue	s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Ketrez (2012), Stachowski (2014), Demir (2017), van Schaaik (2020), Yılmaz (2020), Köylü (2021)	23
mitil		s	Yılmaz (2020)	
mor	purple	s	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Köylü (2021)	23
pak	pure	m	Yılmaz (2020)	3
		s		3
pembe	pink	s	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020), Köylü (2021)	24
perişan	miserable	m	Demir (2017)	
		r	Wedel (1999), Kelepir (2001), G & K (2006), Stachowski (2014), Demir (2017), Yılmaz (2020)	
		s	Demir (2017), Yılmaz (2020)	
pıtırak		s	Yılmaz (2020)	
pis	dirty	m	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
rahat	comfy	p	Özçelik (2012), Yılmaz (2020)	
renkli	colorful	p	Stachowski (2014)	
sade	plain	p	Wedel (1999), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	

sağ	alive	p	Wedel (1999), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
sağlam	healthy	p	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	27
sakin	calm	p	Yılmaz (2020)	
salak	stupid	m		1
		p	Wedel (1999), Yılmaz (2020)	6
sarı	yellow	p	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020), Köylü (2021)	23
sebil		r	G & K (2006)	
sefil	poor	p	Demir (2017)	1
		r	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Stachowski (2014), Demir (2017), Yılmaz (2020)	23
serin	calm	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
sert	stern	m	Swift (1963), Stachowski (2014), Yılmaz (2020)	16
		p	Wedel (1999), Özçelik (2012), Yılmaz (2020), Köylü (2021)	4
sessiz	quiet	m		4
		p	Köylü (2021)	10
seyrek	sparse	p	Özçelik (2012), Yılmaz (2020)	
sıcacık	warm	m	G & K (2006), Özçelik (2012)	
		p	Kelepir (2001), Özçelik (2012), Yılmaz (2020)	
sıcak	hot	m	G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014),	

			Demir (2017), Yılmaz (2020)	19
		p	Özçelik (2012), Stachowski (2014), Demir (2017), Yılmaz (2020)	13
sık	frequent	m	Wedel (1999), Stachowski (2014), Yılmaz (2020)	
sıkı	strict	m	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	19
		p	Özçelik (2012), Yılmaz (2020)	4
sıkkın	depressed	p	Wedel (1999)	
sırsıklam*	--	r	Wedel (1999), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
sıpıç		r	Yılmaz (2020)	
sıska	scrawny	p	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
silik	meek	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
sivri	sharp	p	Wedel (1999), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	26
siyah	black	m	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	29
		p	Özçelik (2012), Stachowski (2014), Yılmaz (2021)	
soğuk	cold	p	G & K (2006), Wedel (1999), Özçelik (2012), Ketrez (2012), Stachowski (2014), Yılmaz (2020)	
soluk	pale	p	Özçelik (2012), Yılmaz (2020)	
sökük	unravelled	p	Özçelik (2012), Yılmaz (2020)	
sulu	watery	p	Wedel (1999), Özçelik (2012), Yılmaz (2020)	19

şekerli	sugary	m		1
		p	Wedel (1999)	9
şirin	pretty	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	8
takır		m	Wedel (1999), Kelepir (2001), Yılmaz (2020)	
		p	Yılmaz (2020)	
talihsiz	unfortunate	p	Köylü (2021)	
tam	complete	s	Stachowski (2014)	
tamam	complete	m		1
		p	Özçelik (2012), Yılmaz (2020)	
		s	Swift (1963), Wedel (1999), Kelepir (2001), G & K 2006, Özçelik (2012), Stachowski (2014), Demir (2017), van Schaaik (2020), Yılmaz (2020)	28
tatlı	sweet	m	Demir (2017)	
		p	Wedel (1999), Kelepir (2001), Özçelik (2012), Demir (2017), Stachowski (2014), Yılmaz (2020)	17
		s	Demir (2017)	
taze	fresh	m	Stachowski (2014)	2
		p	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	30
tek	sole	p	Özçelik (2012)	
		m	Özçelik (2012), Yılmaz (2020)	
tekerlek		p	Özçelik (2012)	
		s	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
temiz	clean	p	Özçelik (2012), Stachowski (2014),	

			Yılmaz (2020)	
		r	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Demir (2017), Yılmaz (2020), Köylü (2021)	22
tengerlek		s	Yılmaz (2020)	
tıkız	dense	m	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		p	Özçelik (2012)	
tok	full	m	Stachowski (2014)	
top	round	r	Wedel (1999), Özçelik (2012), Yılmaz (2020)	
		s	Wedel (1999), Özçelik (2012), Stachowski (2014)	
topaç		r	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		s	Wedel (1999), Kelepir (2001), Özçelik (2012), Yılmaz (2020)	
topan		s	Yılmaz (2020)	
toparlak	round	r	Özçelik (2012)	
		s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Yılmaz (2020)	
tozlu	dusty	p	Özçelik (2012)	
turuncu	orange	p	Wedel (1999), Özçelik (2012), Köylü (2021), Yılmaz (2020), Stachowski (2014), Kelepir (2001)	
tuzlu	salty	m		4
		p	Wedel (1999), Özçelik (2012), Yılmaz (2020), Stachowski (2014), Kelepir (2001)	15
yabancı	foreign	p	Stachowski (2014), Yılmaz (2020)	
yakın	close	p	Özçelik (2012), Yılmaz (2020), Stachowski (2014)	11

		s		6
yakışıklı	handsome	p	Wedel (1999), Yılmaz (2020)	
yalabık	glittering	r	Yılmaz (2020)	
		s	Yılmaz (2020)	
yalıncak		p	Özçelik (2012), Yılmaz (2020)	
yalnız	alone	p	Wedel (1999), G & K (2006), Özçelik (2012), Ketrez (2012), Stachowski (2014), Yılmaz (2020)	25
yamalak		p	Özçelik (2012), Yılmaz (2020)	
		s	Özçelik (2012), Yılmaz (2020)	
yanlış	wrong	p	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
yapışık	attached	p	Özçelik (2012), Yılmaz (2020)	
yarık	broken	p	Swift (1963), Stachowski (2014)	4
		s		1
yassı	flat	m	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	25
		p	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	4
yaş	wet	m	Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	29
		p	Wedel (1999), Yılmaz (2020)	
yaşlı	old	m		1
		p	Wedel (1999), Yılmaz (2020)	13
		s		1
yavaş	slow	p	Wedel (1999), Yılmaz (2020)	
yengil	light	p	Demir (2017)	
yeni	new	p	Wedel (1999), Kelepir (2001),	

			G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020)	
		s	Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
yeşil	green	m m	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), van Schaaik (2020), Yılmaz (2020), Köylü (2021)	30
		p	Özçelik (2012), Stachowski (2014), Yılmaz (2020), Köylü (2021)	2
		s	Köylü (2021)	
yırtık	torn	m		1
		p	Özçelik (2012), Yılmaz (2020)	10
		s	Özçelik (2012), Yılmaz (2020)	3
yoğun	intense	p	Wedel (1999), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
		s	Wedel (1999), Özçelik (2012), Yılmaz (2020)	
yumru	round	p	Özçelik (2012), Yılmaz (2020)	
		s	Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Demir (2017), Yılmaz (2020)	
yumuşak	soft	m	Yılmaz (2020)	7
		p	Özçelik (2012), Yılmaz (2020), Köylü (2021)	5
		r	Yılmaz (2020)	
		s	Wedel (1999), Kelepir (2001), Özçelik (2012), Stachowski (2014), Yılmaz (2020), Köylü (2021)	7
yuvarlak	round	m	Köylü (2021)	1
		p	Özçelik (2012), Demir (2017), Yılmaz (2020), Köylü (2021)	

		s	Wedel (1999), Kelepir (2001), G & K (2006), Ketrez (2012), Özçelik (2012), Stachowski (2014), Demir (2017), Yılmaz (2020), Köylü (2021)	29
yüce	almighty	p	Özçelik (2012), Yılmaz (2020)	
		s	Özçelik (2012), Yılmaz (2020)	
zayıf	weak	p	Swift (1963), Wedel (1999), Kelepir (2001), G & K (2006), Özçelik (2012), Stachowski (2014), Yılmaz (2020)	
zengin	wealthy	p	Wedel (1999), Yılmaz (2020)	11
zor	difficult	p	Stachowski (2014), Yılmaz (2020)	

* This form has no bare root, the reduplicated form is shown here