METATHESIS OF STOP-SIBILANT CLUSTERS IN MODERN HEBREW: A PERCEPTUAL INVESTIGATION

by

Kyle S. Jones

A Thesis Submitted to the Faculty of the

SCHOOL OF MIDDLE EASTERN AND NORTH AFRICAN STUDIES

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF ARTS

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

STATEMENT BY AUTHOR

The thesis titled *Metathesis of Stop-Sibilant Clusters in Modern Hebrew: A Perceptual Investigation* prepared by *Kyle S. Jones* has been submitted in partial fulfillment of requirements for a master's degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

SIGNED: *Kyle S. Jones*

APPROVED BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

_____ July 5, 2016

Samira Farwaneh — Date

Associate Professor of Arabic Linguistics

ACKNOWLEDGEMENTS

I would like to begin by thanking my advisor, Prof. Samira Farwaneh, and the other members of my committee, Prof. Adam Ussishkin and Prof. Mahmoud Azaz. I have had many stimulating discussions with Samira about aspects of Modern Hebrew and Arabic phonology and our mutual skepticism about Semitic roots. Taking her class on Arabic linguistics allowed me to add a new language to my linguistic repertoire, while strengthening my overall understanding of phonology and morphology. She has helped me tremendously, and I am very grateful for her support and encouragement. Adam, likewise, has given me much support, encouraging me to try DMDX and allowing me to conduct my perception experiment in the PsyCoL lab's sound booth. Adam's vast knowledge of Modern Hebrew linguistics has been incredibly useful as well, and his class on Modern Hebrew structure was an important influence on this work.

I would also like to thank Prof. Natasha Warner, who taught me phonetics and allowed me to conduct this research under the IRB of her lab, the Douglass Phonetics Laboratory. I deeply appreciate her assistance with the pilot study and with this larger project.

I am deeply indebted to my Hebrew instructors over the years, including Galia Franco, whose classes sparked my love for Modern Hebrew, Anat and Uri Maimon, whose third and fourth year Modern Hebrew classes were instrumental in improving my Modern Hebrew, and Prof. J. Edward Wright, who taught me the mechanics of Biblical Hebrew grammar. I would also like to thank my Modern Hebrew-speaking friends Tali, who I met in Sacramento, and Talya, who I met in Tucson, who have spoken Hebrew with me and allowed me to ask about finer points of Modern Hebrew grammar and bounce my ideas about the language off them.

Many thanks go to my friend Zoe Kosoff. We encouraged each other to stay on track with our thesis writing, meeting at the library and Caffé Luce to work on our respective theses

for as long as we could stand it during the day. I am also grateful for our crazy adventures at various conferences and when it came time to take a break, and I will miss her company when she goes off to Jordan next year.

I could not have done the statistics required to analyze my data without the help of my friend Renee, who has been my friend since the "good old days" in my first master's program in linguistics at UC Davis. (All errors of implementation and interpretation are, of course, my own). She and Brian have supported me and supplied me with good wine, which is always a welcome occurrence.

Special thanks go out to my parents, Neal and Melanie, and my aunt Melinda as well, who host me in Casa Grande whenever they are down being snowbirds from Lake Tahoe. I would not be where I am without the support of my parents, both financially and emotionally, and I thank them for encouraging me to do academically what made me happy and most interested me. I thank my brother and sister as well, for their encouragement.

I would also like to thank my supervisor at the Bureau of Reclamation, Lisa Lewis, for allowing me to take a few days off here and there to finish up this thesis. She has always encouraged me both at work and in my personal and academic lives, and I am grateful for her support.

Last, but not least, I would like to thank Mike (and the cats too) for being there when I needed moral support.

DEDICATION

In memory of my mother

Charlene J. Jones

1958 - 1998

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In *binyan hitpa'el*, the reflexive and reciprocal verbal conjugation in Modern Hebrew, the /t/ of the /hit-/ prefix categorically metathesizes with a following sibilant (/s/, /z/, /ʃ/, or /t͡s/), giving forms like [histakel] instead of expected forms like *[hitsakel]. It has been theorized that this metathesis may be perceptual, serving to place the /-t-/ in prevocalic position where it can be better perceived by listeners, the direction of metathesis being the more common sibilant + stop sequence in Modern Hebrew (Hume 2004), or that it may be auditory, based on a tendency for the sibilant noise to decouple from the rest of the speech stream, resulting in listener confusion about the place of the sibilant within the word (Blevins & Garrett 2004). Based on data from a speech perception experiment using English speakers, who listened to masked stimuli similar to *hitpa'el* verbs, I argue that Blevins & Garrett (2004)'s account is correct, with English speaking listeners evincing a tendency to misperceive stop + sibilant sequences as sibilant + stop sequences, despite the higher frequency of stop + sibilant sequences in English.

1. Introduction

Until fairly recently, linguists have often thought of metathesis, the switching of sounds within a word, as a sporadic, phonetically unnatural phenomenon, a process un-amenable to standard phonological analyses, connected with speech errors in adults and with the phonological "mistakes" that children make during language acquisition (Hume 2001; Blevins & Garrett 1998). The Neogrammarians, a group of German comparative linguists active in the 1870s, excluded it entirely from their definition of sound change, because metathesis, in their view, was neither "regular" nor "purely phonetically conditioned"; though common historically, they believed that it was due to extra-phonetic factors, like phonological mistakes on the part of speakers (Hale 2003). Later, structuralists like Grammont argued that metathesis occurred to optimize the phonotactics of a word, to take a sequence which was somehow phonotactically irregular or otherwise difficult for speakers to produce and transform it into a sequence which satisfied the particular language's phonotactic requirements (Blevins & Garrett 1998). This had the effect of re-orienting the conversation about metathesis back to the phonetics of the sequences undergoing metathesis, which were thought of as being marked in some way. Metathesis, then, produced less marked structures from marked structures, optimizing the phonetics so that speakers could use less effort in producing the metathesized sequences.

In their pioneering work on phonology, *The Sound Patterns of English*, Chomsky & Halle (1968) thought that their phonological notation for metathesis was too strong, resembling the more complex processes which occur in the syntax, rather than in the phonology. It was only with the turn to Optimality Theory (OT, Prince & Smolensky 1993) and the idea of violable constraints, rather than rewrite rules, that an approach more amenable to metathesis became clear. In OT, the grammar of a given language is thought of as the sum of the interactions

between a set of universal constraints (Prince & Smolensky 1993). These constraints are of two types: 1) Faithfulness constraints, which ensure that the phonological input is faithful to the surface output, militating against phonological changes like deletion, epenthesis, and metathesis, which change the output and differentiate it from the underlying form; and 2) markedness constraints, which militate against marked forms in the output, penalizing violations of a language's phonotactics (Prince & Smolensky 1993). Each language variety has a different ranking of these two types of universal constraints with respect to each other; the constraint ranking determines which possible surface candidate (of an infinite set of possible candidates) wins, the most optimal candidate, which violates the least number of constraints, being the winner (Prince & Smolensky 1993). Importantly, higher ranked constraints can force violations of lower ranked constraints which result in output forms which are unfaithful or marked (Prince & Smolensky 1993).

This last point is important because candidates which violate higher ranked markedness constraints can be ruled out, leaving metathetic forms as the winning candidates, even though they violate faithfulness constraints. In an OT framework, a single constraint, LINEARITY, penalizes any violation of linear order within a word, including metathesis (Horwood 2004). Thus, metathesis can be accounted for by ranking a specific markedness constraint above LINEARITY; Hume (2001) does so in her discussion of metathesis, arguing for a ranking of AVOID X/Y ("Avoid phoneme X in the vicinity of Y," for every possible phoneme) >> LINEARITY. Such a ranking produces metathesis, through avoidance of a particular combination of sounds, in favor of some sequence that is less marked than the sequence being avoided (Hume 2001). This

approach is a continuation of the phonotactic optimization approach[1] that Grammont argued for, formalizing it in a constraint-based phonological theory. Metathesis occurs because of the existence of a dispreferred sequence which can only be repaired through metathesis, because the constraint ranking of the language forces this particular outcome. The OT analysis has an advantage over rule-based analyses because metathesis can be accounted for with a general constraint which does not apply only to metathesis, but which penalizes any kind of linear violation of order.

While this approach to metathesis has been fruitful (and Hume, in particular, has been very productive), it has its own problems. Blevins & Garrett (1998, 2004) argue against the phonetic or phonotactic optimization analysis of metathesis because metathesis can produce output forms which are *more* marked than the input and which are harder for speakers to pronounce. One example they note is the case of glottal stop metathesis in Zoque, a Mixe-Zoquean language spoken in Mexico:

(1) V_1 + ʔV_2 → V_1V_2ʔ (518):

/ʔʌʔwʌ-ʔaŋi/		→ ʔʌʔwaʔŋi	"to that one"
/minʌ-ʔaʔ/	→ minaʔʔ	→ minaʔ	"come now"
/kenu-ʔaʔa/	→ kenwaʔʔa	→ kenwaʔa	"he already looked"
/homi-ʔaŋge/		→ homjaʔŋge	"until tomorrow"

They argue that this metathesis places a glottal stop in word-final and pre-consonantal positions, both of which are worse, more marked positions for glottal stops than their original pre-vocalic positions (Blevins & Garrett 1998). In fact, pre-vocalic or intervocalic position is the optimal place for a stop, because stop perception is totally reliant on external cues, like stop bursts and formant transitions into vowels, which are perceived best before vowels (Hume 2001). It is

[1] This theory is variously referred to as "phonetic optimization" (Blevins & Garrett 2004), "phonotactic optimization" (Blevins & Garrett 1998), and "perceptual optimization" (Hume 2001). These terms will be used interchangeably.

difficult to believe that this is a kind of phonetic optimization, because it does not, in fact, optimize the phonetics of the metathesizing sequences.

Blevins & Garrett (1998, 2004) take a typological approach, emphasizing the different historical pathways of phonetic change that have produced metathesis. In their view, different phonetic processes result in distinct types of metathesis: Long phonetic cues like labialization, for example, can lead to misconstrual of the original place of the long phonetic cue, with listeners reinterpreting the phonological form originally intended by the speaker (Blevins & Garrett 1998, 2004). The result is a metathesized form. Blevins & Garrett (2004) ultimately suggest a four-way typology, based on the phonetic reasons for consonant-consonant (CC) metathesis:

(2) Typology of metathesis (120):

Perceptual metathesis	Elongated phonetic cues
Compensatory metathesis	Stress-induced temporal shifts
Coarticulatory metathesis	CC coarticulation
Auditory metathesis	Auditory-stream decoupling

Perceptual metathesis represents the most common form of metathesis, but, unlike the proponents of the phonetic optimization theory, they restrict the term to only those metatheses which are induced by confusion of the place of elongated phonetic cues. For example, auditory metathesis, which is also widespread, occurs when sibilants metathesize with other consonants, this not being included in the category of perceptual metathesis because Blevins & Garrett (2004) believe there is a different phonetic process that creates this metathesis.

In this thesis, I address the issue of different theories of metathesis in relation to the problem of Modern Hebrew sibilant metathesis, in which a stop /t/ metathesizes with a following sibilant in the reflexive verbal conjugation. Sibilant metathesis is an interesting case because it is disproportionately represented amongst CC metatheses, as Ultan (1978) argues, based on a

largescale typological survey, occurring with a high degree of frequency amongst the world's languages. Sibilant metathesis is also a special case in Blevins & Garrett (2004)'s evolutionary phonology theory of metathesis, because it occurs not due to "elongated phonetic cues," but through a process called "auditory-stream decoupling," whereby the sibilant noise is dissociated from the rest of the sibilant and reinterpreted by listeners in a different position (120). This is an compelling theory that is supported by speech perception research by Bregman (1990), which demonstrates that listeners process speech by dividing the sounds into discrete auditory streams. When faced with a word containing sibilants or other fricatives, listeners segregate the sibilant onto a different auditory stream from the rest of the phonemes because fricative noise is so different from other sounds and is processed in a way similar to non-speech sounds (Bregman 1990; Makashay 2001). When listeners reintegrate these streams in the process of speech perception, they have difficulty determining the original place of the sibilant within the sequence of vowels and consonants, reinterpreting the place of the sibilant within the word (Bregman 1990; Makashay 2001).

Such an explanation is in contradistinction to phonetic optimization approaches to CC metathesis, particularly that of Hume (2004), who argues that all metatheses are perceptual and occur because of indeterminacy in the speech signal. For Hume (2004), indeterminacy may occur because of long duration phonetic cues, which are realized on adjacent syllables, or because of masking by noisier segments. Such indeterminacy leads listeners to reinterpret ambiguous sequences of sounds into the more common attested patterns in their native language (Hume 2004). In the indeterminacy/attestation model, Modern Hebrew sibilant metathesis results from the weakness of the phonetic cues associated with the stop /t/, the sibilant noise obscuring the stop burst, which is its only phonetic cue because of the lack of a following vowel (Hume n.d.;

Hume 2004). The metathesis serves to place /t/ in a pre-vocalic environment, which is the optimal environment for the perception of both its stop burst and the formant transitions into the following vowel. The result is a sibilant + stop sequence, which should be the more common, attested pattern in the language (Hume 2004).

I report here on a speech perception experiment for English speakers using metathesized and un-metathesized Modern Hebrew verbs as stimuli. This experiment is designed to address the following research questions:

1) Which of the two main theories, the indeterminacy/attestation model (Hume 2004) and evolutionary phonology (Blevins & Garrett 2004), adequately describes the facts of Modern Hebrew sibilant metathesis? Is the metathesis perceptual and due to the indeterminacy of the /t/ signal, which is repaired through listeners' language-specific phonotactic knowledge (Hume 2004), or is the metathesis auditory and due to the universal phonetic characteristics of sibilants (Blevins & Garrett 2004)?

2) What do speakers of a different language (English) do when presented with degraded Modern Hebrew stimuli? That is, do they misperceive in the same way as Hebrew speakers might, allowing us to infer that the metathesis occurs because of universal phonetic properties of sibilants?

Because one of the important claims of the phonetic optimization theory is that metathesis creates optimal phonotactic environments from marked phonotactic environments, demonstrating that English speakers misperceive stop + sibilant sequences as sibilant + stop sequences would provide evidence against the indeterminacy/attestation model and for evolutionary phonology, as English and Modern Hebrew have very different phonotactics. Investigating phenomena in one language by testing speakers of another language is a method of linguistic investigation

pioneered by Ohala (1986, 1987), which uncovers the universal phonetic processes which underlie many phonological phenomena. Within this experiment, I compare the speech perception results of English speakers to the surface output of Modern Hebrew speakers, arguing that English speakers misperceive /t/ + sibilant sequences as sibilant + /t/ sequences similar to the way Modern Hebrew speakers metathesize and that the same underlying principle is involved.

This argument does not rely on speech perception measurements from Modern Hebrew speakers performing the same task. Arguably, it makes sense to test Modern Hebrew speakers in the same way as English speakers, but this presupposes that Modern Hebrew speakers' perception is an unknown. But we know that, at some point in the past, Hebrew speakers *did* metathesize a /t/ + sibilant sequence into a sibilant + /t/ sequence, and that this metathesis has continued to the present and even occurs with new sibilant-initial stems. Because of this, I argue that it is sufficient to test English speakers to determine if they do something similar and infer that the same perceptual process applies to both, because phonotactics cannot provide an adequate explanation. Such an outcome is consistent with Blevins & Garrett (2004)'s claim that metathesis involving sibilants is a separate and distinct type of metathesis which occurs due to how the human auditory system processes sibilant consonants. Warner (1998) pursued a similar approach in her investigation of the vowel system of Old Japanese; in her case, she could not test any speakers of Old Japanese because they no longer exist. She could, however, still demonstrate that the set of vowels that she investigated could not have been the proper vowel system of Old Japanese because of the universals of speech perception. I argue similarly in this paper, as the metathesis is also diachronic, in addition to being a synchronic process in the language as it is spoken today.

Importantly, there is another issue with testing Modern Hebrew speakers. Because my experiment uses existing Modern Hebrew verbs which have been manipulated in different ways, Modern Hebrew speakers are always already biased toward the lexical items which exist in the language. I cannot be sure that when a speaker of Modern Hebrew listens to *[hitsaʁek] and answers [histaʁek] that they have actually undergone some process of auditory metathesis or are simply responding with the existing word in Modern Hebrew. Because of this lexical interference from the existing words in the language, different nonsense stimuli would have to be used. For this reason, together with the previously mentioned theoretical issues, I have chosen not to investigate the speech perception of Modern Hebrew speakers, though this is certainly an interesting, but separate, question.

Following this short introduction to the problem in §1, I give a background on the specifics of Modern Hebrew phonology and verbal morphology which are pertinent to the argument. I then discuss the facts of sibilant metathesis in Modern Hebrew in §2, describing the process, three particular instances where it does not occur, but would be expected, and a group of nouns derived from *binyan hitpa'el* verbs which exhibit sibilant metathesis. In §3, I present a review of the literature on metathesis more generally, describing a number of important papers in the study of metathesis, and Modern Hebrew in particular. I also discuss the indeterminacy/attestation model and evolutionary phonology in more detail, contrasting their distinct approaches to metathesis. I then lay out the specifics of my experiment in §4, describing the pilot study which led to this larger study, the stimuli, the participants, and the experimental procedure, which was programmed in DMDX (Forster & Forster 2003), psychological software developed at the University of Arizona. This is followed by an interpretation of the results in §5, using logistic regression and a linear mixed effects regression. I then discuss the results in §6,

arguing that the results demonstrate that sibilant metathesis in Modern Hebrew is due to the way the human auditory system processes sibilants; these results support Blevins & Garrett (2004)'s evolutionary phonology account. I conclude in §7.

1.2 A Note on Modern Hebrew Phonology and Morphology

Modern Hebrew is a member of the Semitic sub-family of the Afroasiatic language family, a large language family which includes ancient Egyptian, the Berber (or Amazigh) languages, the Cushitic languages, the Omotic languages, and the Chadic languages (Greenberg 1963; Hayward 2000). More specifically, Modern Hebrew is the only living member of the Canaanite sub-family of Semitic, though Samaritan Hebrew is still in use as a liturgical language (Hetzron 1987; Lewis et al. 2016). Modern Hebrew is also the most widely spoken revitalized language, with 4.4 million L1 speakers; because of this, the genetic classification of Modern Hebrew is not without debate, as there was an extensive period of time (from the 200s CE until the late 1800s) when Hebrew was not a spoken language (Dekel 2014; Lewis et al. 2016). Indeed, Wexler (1990) argues that Modern Hebrew is a direct descendant of Yiddish, since this was primarily the spoken language of the people who "revived" the language in the late 1800s and early 1900s. This claim has been all but ignored in the linguistic literature, as the theory depends also on the argument that Yiddish is a Slavic language (and that Modern Hebrew, through Yiddish, is also a Slavic language), which is not borne out by the facts (Wexler 1990). Zuckermann (2003, 2009) takes a different approach and argues that Modern Hebrew (which he calls "Israeli") is a hybrid language descended from *both* earlier stages of Hebrew and Yiddish, a significant departure from the model of language descent that sees languages as descendants of a single mother language (Noonan 2010). Despite these arguments questioning the genetic

classification of Modern Hebrew as a descendant of earlier Hebrew, most linguists accept that Modern Hebrew is descended from older Hebrew, albeit with interrupted intergenerational transmission.

Hebrew went through a number of distinct stages in its history, and these are important to understand, as they parallel developments in the history of the Jewish people and are indicative of the unusual development of Modern Hebrew. These stages include Pre-Biblical Hebrew, the earliest stage of the language, which was not significantly different from the Phoenician language; Biblical Hebrew, the language of the Hebrew Bible; Mishnaic Hebrew, the language of the Talmud; Medieval Hebrew, the language of the later exegetical commentators, like Rashi; and Modern Hebrew, from the *Haskalah*, Jewish Enlightenment, in the 1700s to its revitalization as a spoken language (Hetzron 1987). Hebrew was a spoken language in the Biblical and part of the Mishnaic periods, but survived afterward as the language of Judaism and as a lingua franca amongst Jews who did not speak the same vernacular languages (Kutscher 1982). It is likely that Hebrew died out as a spoken language around 200 CE, as there are two distinct layers of Hebrew in the Talmud (Hetzron 1987; Kutscher 1982). Mishnaic Hebrew, the language of the Mishnah, the oldest part of the Talmud (published in 200 CE), shows evidence of being a spoken language used in the daily lives of the Tannaim, the expounders of the Oral Torah, which was eventually gathered together to form the Mishnah (Kutscher 1982). On the other hand, the Gemara, the rabbinical commentary on the Mishnah which forms the other part of the Talmud, is written in Amoraic Hebrew, the language of the Amoraim, who commented on the Oral Torah passed on by the Tannaim, which is clearly a literary language (Kutscher 1982).

Modern Hebrew has departed from earlier Hebrew most significantly in its phonology, which has been significantly influenced by Yiddish. This has included a number of mergers

between phonemes (with spelling sometimes being the only way to determine what the original phoneme was); phonemicization of allophones; and adoption of phonemes from other languages. /ʕ/ and /ʔ/ are now pronounced as /ʔ/ or are not pronounced at all, leading to an incipient sound change which has resulted in long vowels in some cases, while /x/ and /ħ/ have merged into /x/ (Kutscher 1982, 248; Bolozky 1978; Dekel 2014). Similarly, /q/ and /k/ have merged into /k/, and emphatic /tˤ/ and plain /t/ have merged into plain /t/ (Kutscher 1982, 248; Bolozky 1978). These pronunciations are consistent with the Yiddish-influenced Ashkenazi Hebrew[2] pronunciation, though Ashkenazi Hebrew had a spirantized /s/ allophone of /t/, which, under the influence of Sephardi Hebrew, chosen as the prestige pronunciation, is pronounced as /t/ (Kutscher 1982, 248-249).

Table 1: Modern Hebrew Consonants (Laufer 1999; Dekel 2014)

	Bilabial	Labio-dental	Alveolar	Postalveolar	Palatal	Velar	Uvular	Glottal
Stop	p b		t d			k g		ʔ
Nasal	m		n					
Fricative		f v	s z	ʃ ʒ			x~χ ʁ	h
Affricate			t͡s	t͡ʃ d͡ʒ				
Approximant					j			
Lateral			l					

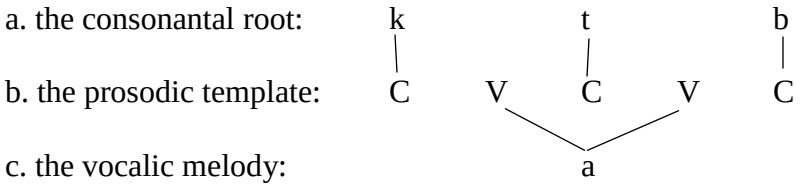
In older stages of Hebrew, the stop consonants /p t k b d g/ (referred to in the Hebrew grammatical tradition by the acronym *BeGeDKeFeT*) spirantized after a vowel, becoming, respectively, /f θ x v ð ɣ/ (Kutscher 1982, 21). The historical reflexes of this phenomenon are still present (except that the variants /θ ð ɣ/ no longer exist), but the phonological process is no longer active in the language (Bolozky 1978; Kutscher 1982). /f x v/ have become phonemes in

[2] Ashkenazi Hebrew is the variety of Hebrew used religiously by Ashkenazim, the descendants of Jews from Western and Eastern Europe. Sephardi Hebrew, on the other hand, represents the variety used by the Sephardim, who were expelled from Spain in 1492 and who fled mainly to North Africa and the Ottoman Empire. "Sephardi" is also a cover term, particularly in Israel, for the descendants of Jews from Arab countries, who are often called Mizrachim ("Easterners") (Pariente 2010).

Modern Hebrew, through internal developments and the influence of other languages, so that minimal pairs now exist: /hitxaveʁ/ “to become a friend of” contrasts with /hitxabeʁ/ “to associate with”; /sixa/ “lubrication” contrasts with /sika/ “pin”; and /safa/ “tongue” contrasts with /sapa/ “sofa” (Kutscher 1982, 249). The loss of gemination, in particular, made the spirantization process opaque and led to its demise. Other differences from earlier stages of Hebrew include the phonemes /ʒ t͡ʃ d͡ʒ/ and the loss of vowel length, so that the vowel system consists of the five vowels /a e i o u/, with the schwa representing a sixth “characteristically unstressed transition vowel” (Chayen 1973, 15; Bolozky 1978; Kutscher 1982). Lexical stress occurs mainly on the final syllable, though some verbal conjugations have penultimate stress, as well as the unusual noun class called the segolates (Bat-El 1989). Only non-ultimate stress is marked in this paper.

Linguists have described Modern Hebrew, in common with other Semitic languages, as displaying a characteristic type of nonconcatenative morphology called root-and-pattern morphology (Prince 1975; McCarthy 1979, 1981). In this conception of Modern Hebrew morphology, words are formed from consonantal roots (generally triconsonantal) and different vocalic melodies which are associated to a particular consonant-vowel (CV) template (McCarthy 1979, 1981). This is distinct from concatenative morphology, as in English, in which words are formed from stems to which are attached affixes in a linear, concatenative derivation. Instead, McCarthy (1979, 1981) posits that Semitic words are composed of separate autosegmental consonant and vowel tiers which are connected to an underlying CV-template; these disparate, non-linear elements are later linearized through a process called tier conflation, which folds all the consonants and vowels into the CV shape required by the CV-template (Ussishkin 1999).

(3) Morphemic tier representation of *katab* (Ussishkin 2005, 171)



In Hebrew grammar, CV-templates correspond to the *binyanim* and *mishkalim*, the verbal and nominal patterns in Modern Hebrew. Verbs in Modern Hebrew fit, traditionally, into one of seven *binyanim*, which are distinguished by their phonological shape and their differing semantics, though some linguists have argued for five *binyanim*, with the two passive *binyanim pu'al* and *huf'al* corresponding simply to the vocalic melody {ua}, which is exchanged for {ie} and {ii} in *pi'el* and *hif'il*, respectively, to passivize them (Berman 1978; Bat-El 2003). The *binyanim* are detailed in Table 2 below, using the example root *k-t-l* "slay," which is the standard example verb. Strictly speaking, there really are no semantics that are necessarily associated with any one *binyan*, except for the marginal *binyanim pu'al* and *huf'al*, which are always the passive forms of *pi'el* and *hif'il* (Berman 1978). However, general claims can be made about the semantics of each *binyan* with the caveat that there are many examples of exceptions. *Binyan pa'al* is generally thought of as being intransitive, with *binyan nif'al* as its passive and *binyan pi'el* as its transitive counterpart. *Binyan hif'il* is causative and *binyan hitpa'el* is generally reflexive and reciprocal. By way of example, the root *k-t-b* has the following verbs in various *binyanim*: [katav] means "write"; [nixtav] means "be written"; [kitev] means "write a lot (literary), engrave"; [hitkatev] means "correspond"; [hixtiv] means "dictate"; and [huxtav] means "be dictated" (Bolozky 2008). It is rare for a root to exist in all *binyanim*; some verbs exist in most of them, some exist in just a few, and some exist in only one.

Table 2: Modern Hebrew *Binyanim* (Adapted from Berman 1978)

ACCEPTED TERM	PA'AL	NIF'AL	PI'EL	PU'AL	HITPA'EL	HIF'IL	HUF'AL
NOTATION	B-1	B-2	B-3	B-4ps	B-5	B-6	B-6ps
Past	katal	niktal	kitel	kutal	hitkatel	hiktil	huktal
Present	kotel	niktal	məkatel	məkutal	mitkatel	maktil	muktal
Future	jiktol	jikatel	jəkatel	jəkutal	jitkatel	jaktil	juktal
Infinitive	liktol	ləhikatel	ləkatel	-	ləhitkatel	ləhaktil	-

Beginning in the late 1980s, some linguists working on Modern Hebrew and other Semitic languages began to question this root-and-pattern model of Semitic morphology, instead suggesting that Semitic morphology is based on stems like the rest of the world's languages (Bat-El 1989). Evidence for this approach comes from denominal verb formation (DVF) in Modern Hebrew (Bat-El 1994; Ussishkin 1999). Bat-El (1994) demonstrates that a root-based account does not work for verbs derived from loanwords with clusters, as clusters are always preserved: The verb [priklet] "to practice law" is derived from the noun [praklit] "lawyer," instead of the expected form *[pirklet] (572). This is important because it indicates that simple root transfer from the noun form to the verb form is not possible; the derived verb also preserves adjacency information as well (Bat-El 1994).

Similarly, Ussishkin (1999) also argues that DVF must reference a stem because of denominative verbs derived from monosyllabic nouns. The output of this kind of DVF is dependent on the vowel of the noun from which it is derived: Verbs of the form $C_1iC_2eC_2$ are derived from nouns with /a/ in the base ([t͡sad] "side" → [t͡sided] "to side"), while verbs of the form C_1ijeC_2 and C_1iveC_2 are derived from nouns with /i/ and /u/ and /u/ and /o/, respectively, ([tik] "file" → [tijek] "to file" and [bul] "stamp" → [bijel] "to stamp"; [sug] "kind, type" → [siveg] "to classify, to sort" and [hon] "capital, wealth" → [hiven] "to capitalize"), with the /u/ alternation being a result of the Obligatory Contour Principle (OCP) (Ussishkin 1999, 405). Even

more interestingly, certain denominative verbs are derived from nouns with affixes: [hitkamt͡sen] "to be stingy" is derived from the noun [kamt͡s-an], which is itself derived from [hitkamet͡s] through the addition of the nominalizing suffix [–an] (Ussishkin 1999, 407). The root of [hitkamt͡sen] cannot be *k-m-t͡s-n*, because the /n/ comes from an affix; if root extraction occurred, only *k-m-t͡s* should have been extracted to form the denominative verb (Ussishkin 1999). The derivation, then, must be based on the entire stem. Ussishkin (1999) makes a similar argument for denominative verbs derived from nouns with /o/, like [koded] "to encode" from [kod] "code," where a vowel in the base persists in the derived output (Ussishkin 1999, 410).

Together these arguments support an account of Modern Hebrew in which the stem is the basic component of a word, not a consonantal root. Ussishkin (2005) extends this analysis to the entire *binyan* system, with *binyan pa'al* as the most basic *binyan* from which all others are derived in an OT output-output correspondence theory framework; this relates surface forms to each other without reference to a more abstract underlying form. Similar analyses have been proposed for Arabic (Ratcliffe 1997, Gafos 2003) as well, and this is the approach I take within this paper, where I deliberately refer to what other linguists might term "roots" as stems.

2. Background

In Modern Hebrew, metathesis occurs categorically in *binyan hitpa'el*, the reflexive, reciprocal, and sometimes inchoative verbal conjugation, which consists of a prefix /hit-/ attached to a verbal stem (Amir Coffin & Bolozky 2005):

(4) /hit-/ + stem: [hitnaʃek] “kiss one another”[3]
[hitlabeʃ] “dress oneself”
[hitʁaɡez] “get excited”
[hitxamem] “become warm”

However, when the stem begins with a sibilant fricative, /s z ʃ t͡s/, the /t/ of the /hit-/ prefix categorically metathesizes with the following sibilant:

(5) [histadeʁ] “arrange onself, get along with”	*[hitsadeʁ]
[hiʃtaɡea] “go crazy”	*[hitʃaɡea]
[hizdaken] “grow old”	*[hidzaken][4]
[hit͡stanen] “catch a cold, cool”	*[hitt͡sanen][5]

Metathesized forms are found in the Hebrew Bible, which means that sibilant metathesis is a very old diachronic phenomenon. Examples of such metathesized forms include [mistateʁ][6] (1 Samuel 23:19), [vajit͡stajaʁu] (Joshua 9:4), [jistaʁgu] (Lamentations 1:14), and [vaʔeʃtameʁ] (Psalms 18:24) (Baden 2010, 40).[7] Sibilant metathesis continues to operate in Modern Hebrew, with new sibilant-initial stems metathesizing as well, as in (3) [histames] “texted one another” (*[hitsames], from the English borrowing “SMS”).

2.2 Non-metathesizing /t/ + sibilant Clusters

It is important to note that this metathesis is a historical process, inherited from earlier forms of Hebrew which continues today, as there are other instances of /t/ + sibilant clusters in

[3] All verbs are given in their third person masculine singular (3MSG) past tense citation forms.
[4] Voicing assimilation occurs with /z/, and complete assimilation can occur with /d/ ([hidaʁdeʁ] ~ [hitdaʁdeʁ] “rolled down, deteriorated”).
[5] The /t/ of the /hit-/ prefix is spelled with the Hebrew letter *tav*, but in metathesized forms with /t͡s/, the *tav* changes to a *tet*. These two letters are pronounced the same in Modern Hebrew, but the spelling change reflects an actual phonetic change in Tiberian Hebrew (Coetzee 1999).
[6] I have transliterated these verbs as they would be pronounced in Modern Hebrew, for ease of exposition.
[7] Interestingly, Baden (2010) also points out that the /t/ of the /hit-/ prefix assimilates to the following sibilant in at least two cases in the Hebrew Bible: Isaiah 1:16 has [hizaku], which Baden (2010) argues is a reflexive and must be in *binyan hitpa’el* (though this is not the traditional translation), and Daniel 12:10 has [vajit͡saʁfu], instead of the expected form [vajit͡staʁfu], which Baden (2010) also believes must be reflexive “they will refine themselves” (40). This may suggest that there was variation between assimilation and sibilant metathesis in earlier forms of Hebrew.

Modern Hebrew which do not undergo metathesis. These clusters, which would otherwise be subject to metathesis because of the immediate adjacency of a /t/ with a sibilant consonant, nevertheless do not metathesize because of historical language change. There are three main groups of words with /t/ + sibilant clusters which do not metathesize: 1) Nouns which begin with the nominalizing prefix /t-/ followed by a sibilant-initial stem; 2) a few rare *binyan hif'il* forms with stems in which the first consonant is /t/ and the second consonant is a sibilant; and 3) borrowings with affricates.

2.2.1 Nouns with the Nominalizing Prefix /t-/

Historically, nouns beginning with the prefix /t-/ followed by a consonant were pronounced with an epenthetic vowel, referred to in the Hebrew grammatical tradition as a "shwa" or "sheva," which broke up what would be an otherwise impermissible sound sequence (Bolozky 1980). Consonant clusters, in fact, only occur due to the action of morphological rules, otherwise never appearing in the language (Prince 1975). This vowel was recorded orthographically by the Masoretes, a group of Torah scribes from the city of Tiberias (hence the name "Tiberian Hebrew" for this stage of Post-Biblical Hebrew), who recorded the pronunciation of the Torah used in the 800s-900s CE so that the proper reading of the Torah would not be lost from future generations (Malone 1993; Prince 1975). Words like /t-guva/[8] "response," /t-muʁa/ "return, compensation," and /t-ʃuka/ "lust" were pronounced as [təguva], [təmuʁa], and [təʃuka]. In the latter noun, /t/ and /ʃ/ were not adjacent, so there was no environment for metathesis to take place.

[8] These are, again, transliterated as they would be spoken in Modern Hebrew.

Under contact with Yiddish, however, these epenthetic vowels were lost in Ashkenazi Hebrew (Bolozky 1994). Yiddish permitted initial consonant clusters, even clusters like /xs/ which are not permitted in Modern Hebrew ([xsidim] in Yiddish being [xasidim] in Modern Hebrew), and this influenced Ashkenazi Hebrew phonologically (Bolozky 1994). Because the first revivers of Modern Hebrew were mainly speakers of Yiddish and acquainted with and influenced by the pronunciation tradition of Ashkenazi Hebrew, Modern Hebrew contains initial consonant clusters as well. Thus, nouns like /t-nua/ "movement" are pronounced [tnua], with no intervening epenthetic vowel. One consequence of this change is that /t-ʃuka/ is now pronounced [t͡ʃuka], with a phonetic affricate (though it is perhaps debatable whether this is phonemically an affricate or not). Vestiges of the earlier pronunciation with an epenthetic vowel do remain in certain clusters: /t-t͡suʁa/ "form" is pronounced both as [t͡suʁa], with complete assimilation of the /t-/ prefix to the affricate /t͡s/, and as [tet͡suʁa], with an epenthetic vowel (Scharzwald 2005).

2.2.2 *Binyan Hif'il* Verbs

Additionally, there are at least two verbs which, when conjugated in *binyan hif'il*, create adjacent /t/ + sibilant clusters which do not undergo metathesis. *Binyan hif'il*, the prototypically causative verbal conjugation, usually has the form hiCCiC, which forces the adjacency of the first and second consonant in the verbal stem. If the first consonant of the verbal stem is /t/ and the second consonant is a sibilant, this results in a /t/ + sibilant cluster, as in the following two verbs:

(6) *Binyan hif'il* verbs with clusters (Hume n.d.):

[hitʃiʃ] "exhaust"
[hitsis] "ferment"[9]

These two verbs in *binyan hif'il* are distinct from the verbs in *binyan hitpa'el* because the /t/ is part of the stem, not an affix; in other words, metathesis only occurs in Modern Hebrew across a morpheme boundary. Importantly, they also have a different stress pattern from *binyan hitpa'el*, which has primary stress on the final syllable, as in *binyan hif'il*, but has secondary stress on the antepenultimate syllable, where the metathesis actually occurs.

2.2.3 Borrowings with Affricates

Similarly, borrowed words with affricates do not undergo metathesis. Words like [t͡ʃips] "French fries," from British English slang "chips," preserve the affricates present in the languages from which Modern Hebrew speakers borrowed the word. Cross-linguistically, there is a tendency to avoid problematic sound sequences in borrowing language phonologies by repairing these sequences with epenthesis, and this might have been expected in Modern Hebrew, considering the historical use of epenthesis to break up impermissible consonant clusters (Smith 2009). (Only final consonant clusters are broken up through the epenthesis of a vowel, as in [filem] "film" from English [fɪlm] and [maʁksizem] "Marxism" from English [maɹksɪzm̩]). As Smith (2009) notes, however, use of epenthesis in loanwords depends on the ranking of faithfulness constraints in an OT model. She argues that there are source-base correspondence constraints specific to loanwords, between a "posited Ls [source language] representation," which, if ranked highly, results in borrowing language forms which are very

[9] It should be noted as well that these are very rare verbs.

similar to the source language (Smith 2009). What this formalism suggests is that there are languages which remain highly faithful to the source language input (and this would be expected in borrowing scenarios where borrowing language speakers have some kind of proficiency in the source language, like with Modern Hebrew) or repair the source language input to conform to the borrowing languages' phonologies, either through epenthesis, the tamest strategy which preserves the source language's phonological characteristics best, or by more radical strategies like deletion. Arguably, the loss of epenthetic vowels in Ashkenazi Hebrew under Yiddish influence opened the door for affricates from other languages to be borrowed, resulting in a situation where affricates are now phonemes of Modern Hebrew, through their adoption from other languages.

Bolozky (1980) also points to findings from Bolozky (1978) which demonstrate that speakers of Modern Hebrew, when asked to produce novel *hitpa'el* verbs from loanwords beginning with a sibilant cluster, like [spoʁt] "sports" or [ʃmal͡ts] "grease, fat," do not metathesize /t/ with the following sibilant. Instead of forms like *[histpoʁtet] and *[hiʃtmal͡tse͡ts], speakers produce the forms [hitspoʁtet] and [hitʃmal͡tse͡ts] (Bolozky 1980, 797). This outcome further supports the idea that Modern Hebrew is a language which preserves the form of loanwords without deleting

2.3 Metathesis Restricted to *Binyan Hitpa'el*

Metathesis does, however, occur in nouns formed from verbs in *binyan hitpa'el*, which serves to illustrate the fact that the process is restricted to *binyan hitpa'el*. The pattern CaCC + [-an] is a productive pattern in Modern Hebrew, forming both adjectives and nouns: [ʁakdan] "dancer" from [ʁakad] "dance" and [salxan] "forgiving, lenient" from [salax] "forgive" (Bolozky

1994, 77). When such a noun is formed from a verb in *binyan hitpa'el*, the resulting noun is metathesized as well, because of its linear derivation from a *hitpa'el* verb: [staklan] "observer" from [mistakel], the *benoni po'el* (present participle) of [histakel] "look, observe" (Bolozky 1994, 79). Thus, the only metathesized forms in Modern Hebrew are connected in some way to *binyan hitpa'el*, any cases where metathesis might have occurred being prevented historically by epenthetic vowels and synchronically by constraints necessitating a high degree of similarity between borrowed words and the forms in the language from which they are borrowed.

3. Literature Review

Metathesis has been, until fairly recently, the neglected "stepchild" of linguistic research (Coetzee 1999). Often swept under the rug as a sporadic, and thus uninteresting, phenomenon, research on metathesis has been largely restricted to the study of patterns of metathesis in historical linguistics, where it is certainly visible and well-known. On the synchronic level of analysis, metathesis has often been reduced to speech errors in adults and acquisition in children. However, despite this superficial analysis of the phenomenon, metathesis has been documented in many different language families across the world, operating synchronically in the grammars of these languages, and it is these crosslinguistic surveys (first Ultan 1978, Hock 1985, then those by Blevins & Garrett 1998, 2004, and Hume 2001, and 2004) that have led to a reassessment of the importance of metathesis in linguistic theory.

3.1 Crosslinguistic Research

The Neogrammarians first proposed the hypothesis that "Sound change takes place according to laws that admit no exception" (Blust 1996). They believed that all sound change

was phonetic and regular, occurring wherever the environment was the same; because metathesis often seems to be an irregular phenomenon, they did not believe that it was due to gradual phonetic change, like any other sound change (Hale 2003). Metathesis and other irregular phonological phenomena, like analogy, were lumped together as extra-phonetic, with Brugmann (1902) claiming that metathesis occurs "when 'the order of sounds and the syllable boundary make for inconvenience'; it causes 'a group of sounds [to be] placed where it is easier for the speaker'" (as cited in Blevins & Garrett 1998, 508). Maurice Grammont, a French Structuralist who studied under Saussure, who, in turn, was influenced by the Neogrammarians, echoed this suggestion, writing that "metathesis 'yields a better syllable structure,' 'safeguards the unity and harmony of a language's sound system in replacing unusual groups by common groups,' 'separates [cluster] types which are or have become unpronounceable in substituting simple types for them,' and 'avoids useless articulatory efforts'" (as cited in Blevins & Garrett 1998, 508-509). These were the beginnings of a systematic phonetic optimization theory of metathesis, which has continued to be influential in the study of metathesis and which was continued in the first largescale crosslinguistic surveys of metathesis.

Later, generative phonologists desired to describe all "phonetic change under one rubric," (Hogg 1977). Kiparsky (1967) and Chomsky & Halle (1968) both have metathesis rules of the following form:

(7) Metathesis rule:

1 2 → 2 1

In these rules, "1" and "2" are phonemes described with distinctive features which switch places with each other. Such rules are important because they demonstrate the desire of generative phonologists to describe and predict metathesis according to rules just like any other

phonological phenomenon; this is a phonological worldview which includes metathesis in the category of sound change, unlike the earlier Neogrammarians and Structuralists (even Bloomfield believed that metathesis was not the result of "gradual drift in performance" (Postal 1968, as cited in Hogg 1977).

Metathesis rules in generative phonology were a result of the recognition that metathesis is often a regular phonological process, studies like Kiparsky (1967), Chomsky & Halle (1968), and Thompson & Thompson (1969) demonstrating this important fact. Thompson & Thompson (1969), in particular, noted that both Rotuman, an Oceanic language, and Clallam, a Salishan language spoken on the Olympic Peninsula in Washington, marked aspectual distinctions with metathesis. In other words, metathesis in both languages is a productive grammatical device, not a random or sporadic phenomenon, as might be expected with slips of the tongue. Rotuman has both long and short forms, with short forms formed from long forms either by apocope of a final vowel or through metathesis of a final CV syllable into a VC syllable:

(8) Rotuman metathesis (Thompson & Thompson 1969; alternative analyses from Biggs 1965):

a. tokiri, tokir "to roll"
   hoto, hot "to jump"
b. seseva, seseav (or sesjav) "erroneous"
   hosa, hoas (or hwas) "flower"
   tiko, tiok (tjok) "flesh"
c. futi, fyt (> *fuit) "to pull"
   famori, famør (> *famoir) "human being"

The distinction between the long and short forms is one of aspect, with the long form ("complete phase") representing a kind of completeness or definiteness (famori ʻea "the people say"), and the short form ("incomplete phase") representing incompleteness or indefiniteness (famør ʻea "some people say") (Churchward 1940, 15, as cited in Thompson & Thompson 1969). Whatever the difference between the aspectual distinctions, Thompson & Thompson (1969) point out that

they are widespread throughout the language. Similarly, Coast Salish languages have a distinction between actual and non-actual forms, one class of which is also marked by metathesis:

(9) Clallam metathesis (Thompson & Thompson (1969):

a. Actual vs. non-actual (no metathesis)
ʔáč̓.t “wipe,” ʔá-ʔ-č̓.t “wiping”
púxʷ.t “blow,” pú-ʔ-xʷ.t “blowing”
məyaʔ.t “kick,” mə-ʔ-yaʔ.t “kicking”
b. Actual vs. non-actual (metathesis)
čkʷú.t “shoot,” čúkʷ.t “shooting”
qi̓qí.t “restrain,” qíq̓.t “restraining”
ƛ̓kʷə́.t “grasp,” ƛ̓ə́kʷ.t “grasping”

Thompson & Thompson (1969) argue that an analysis of this phenomenon as metathesis is superior to analyses with “hypothetical base forms” like *čukʷút, from which a vowel is deleted, or analyses positing “special stress patterns inserting vowels in different positions with relation to root consonants” (217). They argue, specifically, that these kinds of abstractions are unnecessary and obscure important typological patterns, like the similarity between Rotuman and Clallam metathesis, each representing similar kinds of aspectual distinctions (Thompson & Thompson 1969). Importantly, they also call for further investigation of regular metatheses in other languages (Thompson & Thompson 1969).

This view of metathesis as a regular process was not without controversy, however. Webb (1974), in her dissertation on metathesis, analyzes purported metatheses from Standard Arabic, Eskimo, and Kasem (the latter the same example used by Chomsky & Halle (1968) as evidence for metathesis being a common process). In each case, she derives outputs which are not the result of metathesis, but other undisputed phonological processes, like “assimilation, compensation and segmental merger” (Webb 1974, 126). Webb (1974) points out that Chomsky & Halle (1968)’s analysis of Kasem, which establishes their theory of metathesis rules, is

particularly problematic, because it forces them to assume that the underlying representations of other non-metathetic forms are also rearranged, so that the metathesis rule can apply unrestrictedly. Hogg (1977) also expressed qualms about the power of metathesis rules, arguing that metatheses must be marked as exceptions in the grammar, so that these metathesis rules do not overapply, or by creating some more abstract machinery which, while seemingly necessary, is ultimately just a formal device. Her overall conclusion is that metathesis is not a phonological process at all; for her, the only true examples of systematic synchronic metathesis (like Hebrew sibilant metathesis) are morphologically restricted and very rare (Webb 1974). Diachronic metatheses, for her, are likely the result of linguists' lack of knowledge about intermediate stages of languages and the lack of information about suprasegmental factors like prosody (Webb 1974). If such factors were known, Webb (1974) is confident that most examples of metathesis would be reanalyzed as the result of common phonological processes. Webb (1974)'s contrarian views about metathesis did not, however, catch on, and most linguists accept that metathesis is a phonological process, if not exactly common. Crosslinguistic studies, like Ultan (1978) (originally published as a working paper in 1971), have tended to confirm this belief, with the many different examples of possible metatheses strengthening support for its independent existence.

Ultan (1978) is a very important piece of scholarship on metathesis because it represents the first largescale typological survey of patterns of metathesis, work which grew out of Greenberg (1965)'s call for analyzing synchronic language patterns in reference to diachrony (Bickel 2007). Ultan (1978) begins by noting five possible types of formal transposition of elements, which include 1) "inversion of syntactic constituents" as in "He was here vs. Was he here?"; 2) "transposition of syllables," as in the Toba thieves' language, where "tema < mate

‘dead’”; 3) “transposition of sounds,” like “irrevelant < irrelevant” or spoonerisms like “a row of beery wenches” for “a row of weary benches”; 4) “transposition of suprasegmental features,” as in “ímport (n.) vs. impórt (v.)”; and 5) “transposition of phonological features,” as in “Greek thríks (nom. sg.) vs. trikhós (gen. sg.) ‘hair’” (Ultan 1978, 370). Ultimately, he rejects everything but the fifth type as relevant phonological metatheses, though he does note that “While type 4, transposition of suprasegmental features, may be relevant, lack of sufficient examples at present prevents us from including it” (Ultan 1978, 372).

Based on his examination of the data, he provides three hypotheses about metathesis and its operation within a language:

(10) Ultan’s hypotheses (373-374):

a. “Metathesis is a conservative process.”
b. “Metathesis is a recessive process.”
c. “The proneness of different phonetic classes to metathesis tends to stand in direct correlation with a hierarchy of resonance.”

The first hypothesis points out that metathesis tends to be structure-preserving, ensuring that segments which might otherwise be lost are preserved, albeit in a different linear order (Ultan 1978). This is different from other phonological processes, like assimilation, dissimilation, syncope, and epenthesis, which change the phonological output in sometimes drastic ways (Ultan 1978). The second hypothesis refers to metathesis’ tendency to be blocked by other phonological processes; Ultan (1978) notes that this is why metathesis has often been relegated to the realm of “minor sound changes” and why “all-pervasive metathesis” is rarely seen (373). Specifically, he says that “[s]ynchronically, metathesis is considerably limited in scope, although evidently attributable to what must once have been regular phonological causes,” an important statement because it anticipates the later research on metathesis by Hume (1998, 2000, 2001, 2004) and Blevins & Garrett (1998, 2004), which is an attempt to search out these “phonological causes” in

the varying phonetics/phonotactics of different languages (374). The third hypothesis refers to Ultan (1978)'s claim that the more sonorous a segment is, the more likely it is to be involved in a metathesis. In his sample, CV and VV metathesis (together) are very common, with liquids also being very commonly involved in metatheses, a fact that Ultan (1978) calls "proverbial" (375). However, sibilants, "principally s̲," and stops violate this sonority hierarchy of metathesis; Ultan (1978) reports that sibilant metatheses involve stops in all but three of the cases he surveys, an important typological finding (376)

Importantly, he groups the attested metatheses in the survey based on the kind of metathesis which occurs (Ultan 1978). These include: 1) Metathesis involving one or two sonorants; 2) metathesis involving obstruents; 3) quantitative metathesis; and 4) noncontiguous (or long-distance) metathesis (Ultan 1978). The first two types of metathesis are further subdivided based on the segments involved (where W = semivowel, L = liquid, F = non-sibilant fricative, S = sibilant, N = nasal, and P = stop, with representative examples):

(11) Sonorant metathesis (Ultan 1978, 375-378):[10]

a. $V_1 + V_2 > V_2V_1$: Portuguese doesto "affront" < deosto < *denosto
b. $V_1 + W_2 > W_2V_1$: Old Slavic *ej and *ew becoming *ji and *ju
c. V + L > LV: Old Irish túaslucud < túasulcud "opening"
d. L + V > VL: Late Middle English ri > ər before dentals: brid > bird
e. C + V > VC: Rotuman leka [completive] vs. leak, ljak [incompletive] "go"
f. W + C > CW: Zoque j + t > t^j initially and jt^j everywhere else
g. F + L > LF: Cornish whelth "narration" vs. wethlow [plural]
h. P + L > LP: Campadinian Sardinian arbili < aprile "April"
i. P + N > NP: Lithuanian limpù vs. Old Church Slavonic prilĭ(p)na from *leip̲- "smear, stick"

[10] These are written as in Ultan (1978), except where I know the correct IPA transliteration and have substituted it for Ultan (1978)'s notation.

(12) Obstruent metathesis (Ultan 1978, 378-379):

a. P + S > SP: Biblical Hebrew hiʃtammer "he watched himself" < *hit-ʃammer[11]
b. S + P > PS: Indo-European *-sko present metathesizes in Lithuanian before infinitive –ti and future –s suffixes: drẽks "tear (fut.)" vs. drẽskia (3rd pres.), réikʃti "to mean" vs. réiʃkė (3rd pret.)
c. h + P > Ph: Korean jotha < joh-ta "(it) is good," mantha < manh-ta "they are many"
d. C + h > hC: Mandaic Aramaic pāhra "he flies" (cf. Syriac pāraħ)

He also notes a number of different types of long-distance metathesis, with these mainly involving VV metathesis and long-distance metathesis of liquids and nasals and, occasionally, obstruents like fricatives and stops (though metatheses with stops, he claims, are mainly sporadic) (Ultan 1978).

In addition to his classification of metatheses based on the segments metathesizing, Ultan (1978) also investigates the various causes of metathesis, which leads him to a similar four-way typology:

(13) Typology of metathesis causes (Ultan 1978, 383-394):

a. Reduction (imminent and actual)
b. Open syllable canon
c. Phonotactic constraints
d. Various other causes (attraction/repulsion, anticipation of disfavored sound sequences, diphthongization, dissimilation, quantitative equilibrium)

In cases of imminent reduction, metathesis preserves a "reduction-prone segment or feature"; Ultan (1978) points to the "tendency toward apocope of unaccented final vowels" as causing metathesis with a preceding consonant, thereby preserving a final vowel which should otherwise have undergone apocope (384). This occurred in Rotuman if the final vowel was lower than the

[11] Ultan (1978) points out that this metathesis is "extremely common"; the reverse does not occur as frequently: "Thus there seems to be a general preference for clusters of the type sibilant + dental stop over those with the inverse order. This observation leads to a tentative universal: clusters with the order dental (or alveolar) stop + sibilant (i.e. spirant) may metathesize but those with the inverse order do not. The more interesting generalization to the effect that dental + sibilant implies the presence of sibilant + dental may prove to be valid for phonological (as opposed to phonetic) clusters if dental, alveolar and palatal affricates are viewed as unit phonemes" (378-379).

preceding vowel (hwas < hosa “flower,” tjok < tiko “flesh”) or if the final vowel was front or high and “nonidentical” with the preceding vowel, in which case “an umlauted vowel was the result” (fyt < futi “to pull,” høt < hoti “to embark”) (Ultan 1978, 384). In cases of actual reduction, vowel syncope produces metathesis because of stress patterns, as with the Chowchila Yokuts –haḷiẏ consequent adjunctive suffix and the –ilin intensive possessor suffix (Ultan 1978, 386). When suffixed to the absolutive stem, so that the suffix occurs in word-final position, “stress falls on the first syllable of the suffix” (Ultan 1978, 386). But when the suffixes occur before an oblique suffix, “stress is retained on the original penult of the absolutive and the new penult (the second vowel of the disyllabic suffix) is syncopated with metathesis of the two resonants of the resultant inadmissible cluster: xamithaḷiẏ (abs.) vs. xamithaẏla (obj., glottalization lost postconsonantally) ‘scythe,’ pitṭilin (abs.) vs. patṭinli (obj.) ‘one with many body lice’” (Ultan 1978, 387).

Ultan (1978) also points out that the shift to an open syllable canon can create metathesis, as a means of ensuring open syllables in the language. In Old Slavic, due to a number of phonological processes that led to resyllabification, vowel + liquid sequences were metathesized into sequences of non-syllabic liquids and long vowels (prasę “suckling pig,” while the same term in Latin is porcus) (Ultan 1978, 389). Something similar happened in French, with words like fromage “cheese” developing from Latin *formaticum (Ultan 1978, 389). Ultan (1978) also points to a number of “miscellaneous causes,” which are infrequent in his sample and which include attraction (as in Middle English bird, dirt, third < brid, drit, thridde, which he claims may have occurred because of a preference for CVC syllables) and repulsion (bright vs. Old English berht, where the <h> (/x/) apparently repels /r/), which refer to instances where one segment seems to cause another segment to shift; anticipation, whereby “a given string is shifted to a prior

position due to an expected (subconsciously so) difficulty of articulation inherent in the original sequence," as occurs in the Rotuman metathesis case; "the analogical pressure of diphthongization" when a particular diphthong is spreading in a language; dissimilation (as in Spanish milagro < *miraglo "miracle"); and quantitative equilibrium "the preservation of syllable- or word-internal quantity," as may also have occurred with Old Slavic liquid metathesis, because "syllabic quantity could not exceed the length of a long vowel" (Ultan 1978, 393-394).

The most important explanation, in terms of later theories of metathesis, is, however, his "phonotactic constraints" category, as it is a more modern restatement of the earlier phonetic/phonotactic optimization theory of Grammont and others. For Ultan (1978),

> "[m]any metatheses are automatically induced by morphological juxtaposition that results in phonotactically inadmissible sequences. That is, metathesis constitutes a regular morphophonemic process in such instances. Still other metatheses are brought about by the introduction of noncanonic sequences in loan words….Thus the morphophonemic function of metathesis represents a superficial level, usually symptomatic of more general, underlying causes." (390)

Metathesis occurs due to restructuring of sequences which violate a language's phonotactics, either through a morphological process or by adoption of loanwords with phonotactic violations. Thus, in Sopraselva Italian, one finds fartont for Standard Italian frattanto "meanwhile," evincing a "preference for single-consonant initials in syllables" (Ultan 1978, 390). Bagneres de Luchon French also does something similar, as can be seen in burdakin "buskin," where Standard French has brodequin (Ultan 1978, 390). In Spanish, Ultan (1978) argues that r…l > l…r because of a possible preference for the more frequently occurring sequence (as the more common <gr> is substituted for <gl>) (391). Similarly, a final consonant and –n suffix metathesize in Indo-European, likely because of analogy with the more common nasal + stop sequence (Ultan 1978,

392). Such explanations are important because they directly anticipate later arguments from proponents of the phonetic optimization theory, like Hume (1998, 2000, 2001, 2004).

Hock (1985) also adopts a similar approach, arguing that regular metathesis is structural, occurring to satisfy "a general constraint on phonological structure," to dispose of "a cross-linguistically disfavored sequence," or to create "a preferred syllable structure" (529). From the outset, Hock (1985) excludes metatheses that are "clearly morphologically conditioned," like Hebrew sibilant metathesis, because he does not believe these are "regular" (529-530). Likewise, he excludes metatheses with what he refers to as "'epenthesis': the segmentalization of the nonsegmental on- or off-glide of a palatalized or labiovelarized consonant or of a palatal or labiovelar" (Hock 1985, 531). Rather than such cases being instances of metathesis, Hock (1985) argues for an intermediary stage of palatalization or labiovelarization; such would have occurred with, for example, Latin sapiat > *sapya > *sap^{y}a [sa^{y}p^{y}a] > (*)saypa > Spanish sepa "would know" (531). Hock (1985) also excludes metathesis involving /h/ and /ʔ/ for the same reason, because he argues that they are due to segmentalization of aspiration and glottalization. For Hock (1985), examples of true metathesis include Slavic liquid metathesis, where a liquid switches with a preceding vowel, as in Proto-Slavic *mel.ko > mle.ko "milk," and Attic Greek quantitative metathesis, where the lengths of two vowels change places with each other, as in basilḗos > basiléōs "king" (long-short becomes short-long) (533). He argues that each "serves a specific structural purpose," Slavic liquid metathesis creating open syllables while Attic Greek quantitative metathesis "compensates for mora loss under prevocalic shortening" (Hock 1985, 532-533).

The most common "structural purpose" for metathesis is "converting phonologically or perceptually 'marked' structures into more acceptable ones"; this includes metathesis of dental

stop + /l/ (e.g., Old Spanish *tidle > tilde “title” from Latin titulum); stop + homorganic nasal (e.g., Old Spanish *riedna > rienda “rein” from Latin retina); and stop + fricative (e.g., *ksubh- > Lithuanian skùbinti “hasten,” where Sanskrit has kṣubh) (Hock 1985, 533-535). Specifically, Hock (1985) suggests that the common metathesis of stop + fricative sequences into fricative + stop sequences, which otherwise violates the sonority hierarchy (a possible explanation for many metatheses which he wishes to accept), is “perceptually motivated,” explicitly stating that the metathesis could be due to the fact that perceptual cues for stops are better pre-vocalically than they are before another obstruent. This hypothesis is an important precursor to Hume (2000, 2001, 2004)’s work, which makes a similar suggestion about stop + fricative metathesis. Hock (1985) additionally claims that stop + fricative sequences are phonologically marked, regardless of any possible perceptual reason for their tendency to metathesize. He notes that many languages rule out stop + fricative sequences entirely, while other languages, like German, only allow homorganic stop + fricative sequences and others, like Sanskrit and Greek, allow only certain other combinations (Hock 1985).

The introduction of Optimality Theory, which revolutionized phonological theory by exchanging linear rules for non-linear constraints, offered new alternatives for the analysis of metathesis in various languages. Correspondence Theory (CT), a sub-theory of OT, is particularly well-equipped to handle metathesis, as correspondence constraints explicitly formalize relations between input and output, “penaliz[ing] lack of faithfulness between input and output” (Hume 1998). Hume (1998), in her study of two types of metathesis in Leti, argues that metathesis is a distinct phonological process which necessitates the use of a specific correspondence constraint, LINEARITY, which gradiently penalizes violations of linear order from input to output. In other words, output candidate forms can violate LINEARITY more than once,

but the form with the fewest violations of LINEARITY will win (which entails, logically, that there may be optimal forms in some languages with multiple instances of metathesis) (Hume 1998). Leti, an Austronesian language spoken in Timor, displays two distinct types of metathesis, according to Hume: 1) Metathesis ensuring that syllables have onsets and consonant clusters are avoided, these being general "syllable well-formedness conditions" in the language; and 2) metathesis ensuring that "all phrases end in a vowel" (Hume 1998, 147).

The "syllable structure motivated metathesis" is exemplified in the following:

(14) Syllable structure motivated metathesis

a. /kunis + βnutan/	kunsiβnutna	"key + iron"
/ukar + ppalu/	ukrappalu	"finger + bachelor = ring finger"
/ulit + prai/	ultiprai	"skin + prai"
/βuar + spou/	βu:raspou	"mountain + boat = schooner mountain"
/maun + ppuna/	ma:nuppuna	"bird + nest"
/puɔras + liɔra/	pʷɔrsalyɔra	"door + seaside = seaside gate"
/ukar + muani/	ukramwani	"finger + man = middle finger"
b. /ukar + lavan/	ukarlavna	"finger + big = thumb, big toe"
/urun mɔa/	urun mɔa	"breadfruit + Moanese"
/mɛsar lavan/	mɛsar lavna	"teacher, big"
/lout + de/	lout de	"servant, once"

The metathesis serves to ensure that syllable margins are of the type VC.CV, without complex margins (Hume 1998). Indeed, vowel final forms are resyllabified with following complex onsets, as in /koni + mderi/, which is resyllabified as [ko.nim.de.ri] to avoid the complex syllable margin *.mde. (Hume 1998). Complex onsets do, however, occur, but only in "absolute phrase-initial" position: prai "prai," vlakra "crossed," βnutna "iron," mmɛi "table," mlililu "sour," rmori "they live" (Hume 1998, 156-157). However, when the sonority hierarchy is violated (i.e., when the more sonorous consonant precedes the less sonorous consonant in a cluster), the first consonant of a complex cluster is syllabic: ṣpou "kind of boat," ṃderi "Mderi," ḷpia "sago-palm," ṇtutu "he points," ṛpa:ri "they pay" (Hume 1998, 156-157). In OT terms, Hume (1998)

argues that this first metathesis in Leti results from the ranking of LINEARITY below higher-ranked markedness constraints (Hume 1998). These higher-ranked markedness constraints cannot be violated, so metathesis becomes the only optimal alternative; other non-linear analyses using rules are more complex and un-generalizable (Hume 1998). To ensure that complex onsets do not occur, she ranks *COMPLEX (a constraint militating against complex onsets) above CRISP EDGE (a constraint specifying that morphological and syllabic boundaries must be aligned) and LINEARITY (Hume 1998).

(15) Tableau for *COMPLEX >> CRISP EDGE >> LINEARITY (Hume 1998, 155):

		UR: ulit+prai	*COMPLEX	CRISP EDGE	LINEARITY
	a.	u.lit. pra.i	*!		
	b.	ul.ti. pra.i	*!		*
☞	c.	ul.tip.ra.i		*	*

This works, but a possible non-metathesis solution to the problem of complex onsets would be syllabic consonants, as occurs absolute phrase-initially. To prevent this from being the most optimal solution, Hume (1998) ranks the two constraints SONORITY PRINCIPLE ("Between a given sound and the peak are only found sounds of the same, or a higher, sonority class") and CONSPEAK ("Consonants are not peaks") above *COMPLEX and LINEARITY (Hume 1998, 157-158).

(16) Tableau for SP >> CONSPEAK >> *COMPLEX >> LINEARITY (Hume 1998, 158):

		UR: ulit+prai	SP	CONSPEAK	*COMPLEX	LINEARITY
	a.	u.lit. pra.i			*!	
	b.	u.lit. p.ra.i		*!		
☞	c.	ul.tip.ra.i				*

Secondly, Hume (1998) argues that this type of metathesis in Leti also prevents onsetless syllables, as in the surface form uk.ram.wa.ni (*uk.ar.mu.ani) from /ukar + muani/. Avoidance of onsetless syllables is a means of avoiding vowel hiatus and is also resolved by other methods

than metathesis, including realizing a vowel as a secondary articulation on a consonant (/tani + arat/ tan^{y}arta "soil + edge") (Hume 1998, 159). The ranking *COMPLEX >> MAX-V ("No vowel deletion") >> ONSET ("A syllable has an onset") >> *COMPSEG ("a segment may not have more than one place specification" – secondary articulations occur a violation) >> CRISPEDGE >> LINEARITY ensures that an onsetless syllable never surfaces, even at the expense of using metathesis as a repair strategy (Hume 1998, 160-161).

(17) Tableau for *COMPLEX >> MAX-V >> ONSET >> *COMPSEG >> CRISPEDGE >> LIN (Hume 1998, 161):

		UR: ukar+muani	*COMPLEX	MAX-V	ONSET	*COMPSEG	CRISPEDGE	LIN
	a.	ukar muani			*!			
	b.	ukar mwani	*!					
	c.	ukar mani		*!				
	d.	ukar m^{w}ani				*!		
☞	e.	ukram wani					*	*

Thirdly, syllable structure motivated metathesis also occurs with compensatory lengthening effects, which occur when a vowel from a VVC sequence is deleted or moved (Hume 1998). Forms like βu:raspou from /βuar spou/ illustrate this process. This occurs to compensate for a loss of a mora, a measure of syllable weight, which is governed by the constraint DEP-μ, which specifies that the input and output must contain the same number of morae (Hume 1998). Mora deletion violates MAX-μ, and the candidate with metathesis wins (Hume 1998).

(18) Tableau for *COMPLEX >> MAX-μ >> DEP-μ >> LINEARITY (Hume 1999, 163):

		UR: βuar+spou	*COMPLEX	MAX-μ	DEP-μ	LINEARITY
	a.	βu.ar. spo.u	*!			
	b.	βu.ras.po.u		*!		
☞	c.	βu:.ras.po.u			*!	*

Finally, phrase-final metathesis, which occurs with words ending in consonants when they occur phrase-finally (urnu phrase-finally vs. urun mɔa phrase-internally), is fueled by alignment constraint, ALIGN-PHRASE, which ensures that the right edge of a phrase is lined up with a vowel.

(19) Tableau for ALIGN-PHRASE >> LINEARITY (Hume 1998, 164):

		UR: kunis##	ALIGN-PHRASE	LINEARITY
	a.	kunis	*!	
☞	b.	kunsi		*

Hume (1998) provides a unified account of two different types of metathesis in Leti, demonstrating that metathesis occurs because of higher ranked well-formedness constraints in the language. Because these higher ranked constraints, which operate generally within the language, cannot be violated, the metathetic candidates end up being the most optimal candidates. This is, in effect, a formal analysis along the lines of earlier suggestions about phonetic/phonotactic optimization by Grammont and others.

However, Blevins & Garrett (1998), in their survey of CV metathesis, reject this phonotactic optimization analysis, arguing that CV metathesis "has a distinct typological profile which follows quite naturally from the traditional view that sound change arises from the reinterpretation of phonetically ambiguous surface forms" (509). This amounts to an extension of the Neogrammarian's claim about the gradual nature of sound change to metathesis, which the Neogrammarians themselves never linked to their own hypothesis about sound change, because they believed the process was not regular. For Blevins & Garrett (1998), CV metathesis falls into two main categories, based on their diachronic origins: 1) Perceptual metathesis, which occurs with phonetic features of long duration; and 2) compensatory metathesis, which occurs due to perseveratory coarticulation. They also discuss what they call pseudometathesis, and they

include the Leti examples analyzed by Hume (1998) in this latter category (Blevins & Garrett 1998).

Perceptual metathesis is characterized by "segments with certain features [that] jump over adjacent segments"; these long duration phonetic features include laterality, rhoticity, aspiration, glottalization, pharyngealization, palatalization, and labialization (Blevins & Garrett 1998, 509). Each of these features is realized across the VC or CV domain, and, because of their long duration, can cause listeners confusion as to the original place of the feature within the word (Blevins & Garrett 1998). For example, /CVɦC/ can be reinterpreted by listeners as /CɦVC/ if /CVɦC/ is realized as [CV̤C] (Blevins & Garrett 1998, 511). Listeners cannot determine which of the two possible consonant candidates is causing the breathiness in the vowel, and they may reinterpret the breathiness as originating with the other consonant in a CVC string. Indeed, Blevins & Garrett (1998) describe three diachronic stages in perceptual metathesis:

(20) Stages of perceptual metathesis (Blevins & Garrett 1998, 512):

1) Phonological analysis of a string into VC/CV sequences
2) "[F]eatures of the consonant shift or spread across an adjacent vowel," leading to multiple association
3) The "metathesizing consonant is reanalyzed as originating on the other (nonhistorical) side of the vowel in question"

Blevins & Garrett (1998) point to laryngeal metathesis in Cayuga as a possible representation of these steps. In Cayuga, /h/ metathesizes with a preceding vowel in odd-numbered non-final syllables (which are unaccented), as in /kahwistaʔeks/ → [k^{h}ą̣wísdʔaes] "it strikes, chimes (a clock)," /akekahaʔ/ → [agékhaaʔ] "my eye," and /koʔnikõhaʔ/ → [gʔoníkhwaʔ] "her mind" (Blevins & Garrett 1998, 510). The metathesis does not occur, however, in forms like /akahwistáʔek/ → [agahwisdáʔek] "it struck, chimed" (Blevins & Garrett 1998, 510). They hypothesize that "the nuclei of metrically weak syllables were shortened

phonetically," leading to a situation where the sequence /Vh/ was phonetically [V̥], a devoiced vowel (Blevins & Garrett 1998, 511). /CVh/ was [CVV̥] phonetically, and this eventually became CV̥, which was interpreted by listeners as /ChV/ rather than /CVh/, resulting in metathesis (Blevins & Garrett 1998, 511). They believe this also occurred with glottal stops, with the sequence /CVʔ/ being realized phonetically as [CVV̰], with creaky voice (Blevins & Garrett 1998, 511). In an earlier hypothetical historical stage, aspiration and glottalization were associated with an entire VC/CV string; later, as in contemporary Cayuga, this aspiration spread across a nearby vowel (Blevins & Garrett 1998). Finally, Cherokee, a related Iriquoian language, represents the final stage, having reinterpreted "postvocalic laryngeals…as prevocalic" (Blevins & Garrett 1998, 512). Thus, a consonant associated with a long duration phonetic feature hops over an adjacent vowel.

Blevins & Garrett (1998) argue that this account of perceptual metathesis has three advantages over the phonetic optimization analysis. They note that these metatheses are symmetrical; where one finds VC → CV metathesis, one also finds CV → VC in a different language, because the long phonetic features can be reinterpreted in either direction (Blevins & Garrett 1998). This is a blow to the idea that these metatheses are due to universal markedness, that one direction is somehow more optimal than another, based on independent phonological, perceptual, and articulatory evidence (Blevins & Garrett 1998). These metatheses also involve phonetic factors which are easily accounted for in a perceptual account, but which cannot be accounted for from a phonotactic viewpoint. In Latin, a change from LV > LV before a non-coronal is theorized to have occurred, with word like *dlukis > dulkis "sweet" and *plumo: > pulmo: "lung" metathesizing, while words like pluteus "shed" did not (Blevins & Garrett 1998, 516). Blevins & Garrett (1998) note that if /l/ was velarized before /u/, its formant structure

would be very similar to /w/, and listeners could easily misperceive this sequence and assume that the laterality was postvocalic, rather than prevocalic (518). This metathesis did not occur before coronals because coronals have a tendency to front preceding vowels, which would make the formants of /lˠ/ and /u/ sufficiently distinct from one another that they were not confused, thereby preventing metathesis (Blevins & Garrett 1998). A phonotactic approach is, however, forced to make absurd claims about one sequencing being better than the other.

Also, Blevins & Garrett (1998) argue that phonotactic explanations of metathesis make incorrect predictions about extant metathesis patterns. For example, VC > CV metathesis involving obstruents does not occur, though, in theory, one combination might be phonotactically more appropriate than another (Blevins & Garrett 1998). In Blevins & Garrett (1988)'s approach, VC > CV metathesis with obstruents does not occur because there are no long phonetic features which can migrate from their original place. Their approach to CV metathesis adequately constrains possible patterns of metathesis and easily accounts for why metathesis occurs in the first place. They also claim that arguments about the structure-preserving nature of metathesis are specious because children, in the process of language acquisition, have no access to the overarching patterns common in their native languages. Similarly, proponents of phonetic optimality, as the force driving metathesis as a repair strategy of non-optimal segment sequences, are hard-pressed to account for such non-optimal metatheses as Zoque glottal metathesis, which takes an optimal /VʔV/ (as a stop's external phonetic cues are best perceived before a vowel) and creates a non-optimal [VVʔC] or [VVʔ]$_{W}$ sequence.

Blevins & Garrett (1998) also propose a perceptual account of compensatory metathesis, which they theorize is the result of perseveratory coarticulation. Compensatory metathesis can occur at either the left or right edge of a word after a stressed vowel through anticipation of a

following unstressed vowel; anticipation of the unstressed vowel creates a copy vowel on the other side of the adjacent consonant, and, through eventual final vowel loss, metathesis takes place:

(21) Stages of compensatory metathesis (Blevins & Garrett 1998, 527):

Right edge: …$\acute{V}_1C\check{V}_2$] > …$\acute{V}_1\check{V}_2C\check{V}_2$] > …$\acute{V}_1\check{V}_2$C]
Left edge: [$V_1C\acute{V}_2$… > [$\check{V}_1C\check{V}_1\acute{V}_2$… > [$C\check{V}_1\acute{V}_2$…

Compensatory metathesis can only occur if a number of language characteristics are first met: They must have "penultimate stress, small vowel inventories, relatively steady-state (as opposed to diphthongal) vowels, and the absence of consonant clusters, long consonants, and consonants with secondary articulations" (Blevins & Garrett 1998, 534). This explains why the process is not more common typologically, despite the fact that it occurs because of "three fairly common phonetic events: coarticulation or gradual temporal realignment in VCV sequences; tonic length; and peripheral vowel reduction and loss" (Blevins & Garrett 1998, 528, 534).

Evidence for this process comes from languages like Rotuman (previously cited above) and Kwara'ae, an Austronesian language spoken in the Solomon Islands. In Kwara'ae, normal forms, used in everyday speech, differ from citation forms, which are used in careful speech and in caretaker speech, call-out routines, and songs (Blevins & Garrett 1998). The citation forms preserve final vowels, penultimate stress with alternating leftward syllables, and "pure vowels [i u e o a] in all positions" (Blevins & Garrett 1998, 529). Kwara'ae normal forms, on the other hand, have either a final devoiced vowel or have lost their final vowel entirely, depending on the quality of the vowel and the following consonant: Normal forms have devoiced vowels if the preceding vowel is of either greater or equal sonority, ási > ʌisi̥ "sea," or if the preceding vowel is identical, óso > o·so̥ "lie" (Blevins & Garrett 1998, 530). Normal forms have lost their final vowels after nasals, as in lúma > luəm "house," and non-high vowels following high vowels are

also lost, as in *asíla > asíələ̥ > asiəl “salty, sweet” (Blevins & Garrett 1998, 530). A following nasal facilitates this vowel loss because it is difficult to perceive a voiceless vowel after nasal murmur (Blevins & Garrett 1998). Blevins & Garrett (1998) argue that the other instance of vowel loss is due to centralizing and laxing of non-high vowels following a high vowel; these voiceless vowels were shorter and thus more susceptible to loss. Similar cases of compensatory metathesis with copy vowels exist in Ririo, another Oceanic Austronesian language, and evidence for similar stages exists in Pama-Nyungan, a group of Australian languages.

Finally, Blevins & Garrett (1998) propose that the metathesis in Leti, which Hume (1998) analyzes, is actually a case of pseudometathesis, a seeming case of metathesis which arises from different historical origins than the two previously mentioned varieties of CV metathesis. They argue that pseudometathesis is very often the result of epenthesis and subsequent deletion, rather than any true metathesis process due to perceptual factors at play (Blevins & Garrett 1998). As previously mentioned, Leti has two distinct types of metathesis: 1) An internal metathesis involving stem alternations (βu:ra [phrase-finally] ~ βuar [phrase medially before a CV syllable] ~ βu:r [phrase-medially before a V “mountain”]), and 2) an external metathesis (/asu + mermera/ → asmwɛrmɛra “red dog”), which they argue is the result of “resegmentation of elongated palatal and labial phases” (Blevins & Garrett 1998, 542). The first “metathesis” is really the product of three distinct stages of language change. First, vowels were epenthesized after final consonants, a sound change which is widespread in Malayo-Polynesian languages: *ánin > *áninə “wind” (Blevins & Garrett 1998, 542). Final consonants were released, and this release was interpreted by speakers as a reduced vowel (Blevins & Garrett 1998). Later, “unstressed phrase-medial vowel loss in open syllables” occurred phrase-finally, and, in an instance of perceptual metathesis, the qualities of the preceding vowel colored the following shwa vowel,

causing it to be reinterpreted as the preceding vowel: *áninə > *ánni (Blevins & Garrett 1998, 543). Phrase-medially, high vowels before CV syllables, where the second vowel was nonhigh, were reinterpreted by speakers as /j/ and /w/, due to the effects of palatalization and labialization on the following consonant: *βáru # CV- > βar # CwV- "new" (Blevins & Garrett 1998, 543). Non-high vowels in the same environments were simply deleted: *áninə # CV- > ánin # CV- (Blevins & Garrett 1998, 544). Finally, schwa shifted to /a/: *βúlnə > *βúlna (> βúlla) "moon" (Blevins & Garrett 1998, 544). So, while the outcome of these sound changes looks like the familiar types of metathesis that Blevins & Garrett (1998) catalogue, they occurred because of different historical reasons, which Hume (1998) does not and cannot account for. Importantly, the different historical sound changes that Blevins & Garrett (1998) postulate are founded on very well-reconstructed proto-forms, providing a more convincing account of Leti metathesis which harmonizes with diachronic research on its ancestors.

This latter point is crucial to an understanding of the differences between phonetic/phonotactic optimization and Blevins & Garrett (1998)'s diachronic approach to synchronic metathesis. Blevins & Garrett (1998) explain that perceptual metathesis is synchronically analyzable, due to phonetic factors, but compensatory metathesis and pseudometathesis are "problematic," leading to usage of "parochial constraints," particularly in OT; "unworkable representational attempts to eliminate metathesis rules of the form AB → BA from the grammar"; and abstract representations which are unreasonable and not falsifiable, particularly as in Government Phonology (549). A diachronic analysis of these phenomena is natural: "[S]ound changes result from misinterpretation: X changes to Y when X is misinterpreted—and phonologically internalized, and therefore later produced—as Y" (550). Such a "model of phonological diachrony," they argue, is "more economical" than the phonetic

optimization alternative, which must assume that something like "perceptual ease" is incorporated in the grammatical knowledge of the speaker (Blevins & Garrett 1998, 550). Otherwise, there could be no sense of perceptual ease as the "GOAL of the change"; Blevins & Garrett (1998) do not have to make this assumption, noting, instead, that, "We assume only that what is hard to perceive is sometimes not perceived, and that misperception leads to misinterpretation and change" (550).

The use of parochial constraints, in particular, is a common method of explaining metathesis in OT; Hume (1999, 2001) says explicitly that metathesis occurs because of the ranking AVOIDX/Y ("Avoid X in the vicinity of Y") >> LINEARITY. AVOIDX/Y is supposed to penalize non-optimal sequences of sounds, and Hume (2001) says that it should encompass every possible sound combination in any language. This introduces a huge possible number of constraints, most of which would not operate in any one language, and it creates further questions about child language acquisition and learnability.

Hume (2004)'s later work on the indeterminacy/attestation model of metathesis sidesteps the issue of the formal representation of metathesis in phonological theory by focusing on a "unified, explanatory account of why metathesis occurs, why it favors certain sound combinations, and why we obtain the output that we do," though she focuses mainly on CC metathesis (204). Her functional account addresses three important observations about metathesis, namely that: 1) The metathesizing sounds differ across languages (one language, for example, has XY changing to YX, while another has YX changing to XY); 2) some sequences only metathesize in one direction, while others metathesize in both directions; and 3) metathesis often improves phonetic cues associated with one or both of the metathesizing sounds (Hume 2004). Her complex account relies on previous psycholinguistic and language acquisition

research to claim that listeners reinterpret indeterminate (or ambiguous, in the terminology of other work on metathesis) speech signals into the more common sound sequences in their languages by using their experience of their native languages and knowledge of language usage (Hume 2004). This explains why metathesis occurs (indeterminacy of the speech signal); why metathesis creates certain sound sequences and not others (only attested and more common sequences in the language); and why metathesis appears to optimize the phonetics of the speech sounds involved (sequences with better cues are more stable and more common and persist within languages, while those with weaker cues tend to be lost).

According to Hume (2004), both a listener's language experience and knowledge of language usage influences how speech is perceived and processed. In other words, for Hume (2004), speech perception is not a universal process, but is one which is language-specific: "[G]iven the inescapable influence on one's language knowledge, a sequence of sounds with identical phonetic cues may be parsed differently by different listeners (of different languages or even the same language)" (211). This is in contrast to the more Ohalian approach that Blevins & Garrett (1998) take, arguing that all sound change is due to misperception, which, in turn, is due to universal phonetic processes which do not change from language to language. Hume (2004) argues for the important role of language experience and usage using a number of studies from child language acquisition and psycholinguistic speech processing experiments. She points to studies on child language acquisition which demonstrate babies' ability to distinguish their native languages from other unfamiliar languages and their ability to distinguish familiar from unfamiliar words (Werker & Tees 1999, as cited in Hume 2004). In fact, very young babies are able to discriminate between native and nonnative sound contrasts, with this ability diminishing by the age of ten to twelve months (Hume 2004). This diminished capacity to discriminate

phonetic differences which are used by other languages explains why people who learn a second language are worse at distinguishing the sounds of the second language than they are at distinguishing sounds in their native languages (Hume 2004). For Hume (2004), this demonstrates the important role of language experience in speech processing.

Psycholinguistic experiments have also demonstrated speakers' awareness of their native language's sounds, phonotactics, patterns of contrast, and syllable structure (Hume 2004). Dutch and Japanese speakers, for example, use nasal place cues in different ways, due to differences in the phonological systems of each of the languages (Hume 2004). Japanese nasals are always homorganic with a following consonant, so that Japanese speakers can reliably use nasal place cues to infer information about a following consonant; Dutch, on the other hand, does not have this straightforward information, so Dutch speakers ignore nasal place information in deducing the place of a following consonant (Hume 2004). Similarly, knowledge of language usage is important in speech processing as well, with more frequent lexical items being accessed sooner and easier than less frequent items (Hume 2004). Indeed, many studies have found that nonsense words which consist of legal sound sequences in a language are more acceptable than those which violate phonotactics (Hume 2004).

Both of these factors, language experience and language use, are important with respect to the indeterminacy of the speech signal: Listeners reinterpret indeterminate sequences of sounds according to their language experience and their knowledge of language use, the occurring patterns of metathesis representing the more common, attested sequences in the language (Hume 2004). Hume (2004) defines "indeterminacy of order" as "a state in which there is insufficient information concerning the linear ordering of the elements involved," this being

"…a function of two factors:

a. the listener's experience with the elements involved (e.g. sounds, sound sequences, morphemes, words, etc.);
b. the quality of information occurring in the speech signal (e.g. the types of sounds involved, the context in which the sounds occur, the phonetic cues available, etc.)." (216)

The specific interaction between these two factors creates an environment ripe for metathesis.

But what causes the indeterminacy that leads to metathesis of uncommon sound sequences in the first place? Hume (2004) argues that there are two different types of sounds involved in metathesis: Those with "diminished perceptual salience" and those with "a lack of temporal resolution" (217, 220). Effectively, "diminished perceptual salience involve[s] either similar sounds and/or those where the phonetic cues to the identification of at least one of the sounds is masked" (Hume 2004, 217). She notes that 35% of the 37 cases in her study ("Metathesis in Language," Hume n.d.) involve sounds that are acoustically similar to each other, and one third of all CC metatheses involve stops, whose cues are frequently masked by other sounds (Hume 2004). Masking occurs specifically because of the particular phonetic cues of the sounds involved, some sounds, like stops, relying almost entirely on external cues, which can be obscured by other sounds (Hume 2004). Regarding stops, Hume (2004) notes that the silent period of a stop, during its closure, is the only internal phonetic cue which can be used to identify a sound as a stop:

(22) Perceptual cues to obstruent stops (Hume 2004, 214):

TYPE	CUE	SEGMENT-INTERNAL OR CONTEXTUAL CUE
manner:	silence	internal
	release burst	contextual: consonant release
	transition duration	contextual: VC, CV transitions
place:	F2 transition	contextual: VC, CV transitions
	burst spectrum	contextual: consonant release

In contrast, fricatives have many more robust internal cues, and this contrast between obstruent stops and fricatives explains why stop/fricative combinations metathesize so frequently: The

fricative consistently masks the perceptual cues to recovery of stop place and manner (Hume 2004).

(23) Perceptual cues to fricatives (Hume 2004, 215):

TYPE	CUE	SEGMENT-INTERNAL OR CONTEXTUAL CUE
manner:	frication noise	internal
	noise duration	internal
place:	frication spectrum	internal
	frication amplitude	internal
	F2 transition	contextual: VC, CV transitions

Both acoustic similarity and masking can occur at the same time, as Hume (2004) argues is the case with Modern Hebrew metathesis, which involves homorganic sounds (coronal sounds /t s z ʃ t͡s/) and masking by sibilants of the internal phonetic cues of /t/.

A lack of temporal resolution due to "'stretched out' features" also causes listeners to misinterpret sequences of sounds (Ohala 1993, as cited in Hume 2004, 219). The time domain of such features extends over adjacent sounds, leading to acoustic overlap and an inability to determine the order intended by the speaker, as listeners cannot determine where one sound begins and another ends (Hume 2004). Glottals, liquids, and glides all have these "stretched out" phonetic features that are realized across long time domains, and Hume (2004) argues that these sounds are found in a large number of cases of CV and CC metathesis because of this. In a footnote, she speculates that listeners' inability to recover temporal order may be due to "auditory-stream segregation," a theory of speech perception proposed by Bregman (1990). According to the theory of auditory scene analysis, "separate auditory continua or streams are created among similar auditory cues and remain perceptually separated without temporal cross-linking" (Hume 2004, 220). In other words, listeners perceive dissimilar sounds on distinct auditory streams, which are not connected temporally. When it comes times for listeners to process this speech, they then have difficulty reintegrating the auditory streams, resulting in

errors in temporal order. Interestingly, Blevins & Garrett (2004) use this same reasoning to argue for a separate category of auditory metathesis, involving fricatives, separate from metathesis involving long phonetic cues, which they term perceptual metathesis.

Importantly, Hume (2004) argues that metathesis always creates an output which is attested in the language, these metathetic sequences usually being the more common sequences in the language, due to listeners' knowledge of and expectations about their own languages. This idea is "consistent to an extent with earlier proposals suggesting that by metathesis, uncommon language structures are replaced by more common ones (see e.g. Grammont 1933, Ultan 1978, Hock 1985)"; however, Hume (2004) contends that this is language-specific, not universal. Specifically, this last claim explains why languages differ in the direction of metathesis. An order of sounds that is common in one language may be uncommon in another, causing different directional biases for listeners (Hume 2004). In fact, "in principle, any order of two segments is a potential output of metathesis, provided that the reordered sequence forms an attested structure in the language" (Hume 2004, 221). In the case of there being only one attested sequence, the listener will choose that sequence in a case of indeterminacy; if there are two possible sequences, then listeners choose the "most robust sequence, that is, the one with the highest frequency" (Hume 2004, 222). For Hume (2004), this means that Modern Hebrew metathesis creates a sibilant + /t/ sequence because /t/ + sibilant is less common, only occurring tautomorphemically ([hi-tsis "he fermented") or in "nonverbal forms" ([t-ʃuva] "reply") (Hume 2004, 222). Similarly, speakers of American English often pronounce "chipotle" as [t͡ʃɪ.poʊl.ti] (instead of [t͡ʃɪ.poʊt.lej]), with metathesis of the /tl/ sequence to an /lt/ sequence, because, Hume (2004) claims, the /lt/ sequence is more common in English. In both cases the resulting sequences are

also better perceptually, something that, for Hume (2004), occurs for free, this phonetic optimization not being the teleological goal of metathesis.

However, metathesis does not have to create phonetically optimized sequences because "speech processing is not universal" (223). Thus, in Mutsun, a Costanoan language formerly spoken in California, the nominal thematic plural suffix [-mak] alternates with [-kma] postvocalically; the locative suffix [-tak] also alternates postvocalically with [-tka]:

(24) Mutsun metathesis (Hume 2004, 224):

ruːk	ruːk-mak	"string"
ʔinnis	ʔinnis-mak	"son"
rukka	rukka-kma	"house"
tʃiri	tʃiri-kma	"paternal aunt"
ʔurkan	ʔurkan-tak	"mortar"
loːt	loːt-tak	"mud"
siː	siː-tka	"water"
paːrani	paːrani-tka	"hill"

What is interesting about this alternation is that the expected alternant of [-mak] (based on the [-tak]~[-tka] alternation) should be *[-mka], which is non-occurring; instead, there is long-distance metathesis of the /k/ (Hume 2004). Hume (2004) explains this unexpected metathesis pattern as a result of the overall patterns in the language, which she theorizes listeners use to process speech. The sequence /mk/ is uncommon, while the sequence /km/ is very common in Mutsun outside of this particular alternation (Hume 2004). The metathesis from /mk/ to /km/ is unexpected from an approach that relies purely on phonetic cues, as the sequence */-mkV/ is more favorable from a phonetic optimization perspective, as stop cues can be best perceived

before vowels (Hume 2004). For Hume (2004), this is further evidence that speakers utilize language knowledge to perceive indeterminate sound sequences.[12]

Blevins & Garrett (2004) continue their diachronic approach to metathesis begun in Blevins & Garrett (1998) by expanding their typology of metathesis to include both CV and CC metathesis. They situate their work as diachronic and nonfunctionalist in a typology of explanations of sound change:

(25) Typology of sound change theories (Blevins & Garrett 2004, 119):

a. Synchronic + nonfunctionalist
b. Synchronic + functionalist (e.g. Flemming 1996; Hume 1997, 2001; Boersma 1998; Steriade 2001)
c. Diachronic + functionalist (e.g. Grammont 1950; Vennemann 1988)
d. Diachronic + nonfunctionalist (e.g. Ohala 1974, 1981, 1993; Blevins and Garrett 1998)

This is a view of sound change which highlights the role of phonetics in motivating sound change, while rejecting the idea that the goal of metathesis is phonetic optimization. Blevins & Garrett (2004) argue, as they do in Blevins & Garrett (1998), that all sound change is the result of misperception by listeners, with any phonetic optimization occurring simply as a byproduct of metathesis, rather than its goal: "…change [often] *results* in perceptual optimisation, for the natural reason that what is harder to hear is sometimes not heard, but our account does not invoke perceptual optimisation as a mechanism or cause of the change" (133).

Expanding on their two-way typology of CV metathesis, Blevins & Garrett (2004) argue for a four-way typology of metathesis, with corresponding phonetic features which lead to each type:

[12] Alternatively, Butler (2013) argues that this metathesis in Mutsun has no phonetic basis, resulting instead from "strong prosodic constraints…on the penultimate syllables of words, which are preferentially heavy (suffix metathesis)" (319). Such an explanation undermines Hume (2004)'s arguments.

(26) Typology of metathesis (Blevins & Garrett 2004, 120):

Perceptual metathesis	Elongated phonetic cues
Compensatory metathesis	Stress-induced temporal shifts
Coarticulatory metathesis	CC coarticulation
Auditory metathesis	Auditory-stream decoupling

Blevins & Garrett (1998) previously discussed perceptual and compensatory metathesis in the context of CV metathesis, but they add coarticulatory and auditory metathesis to their metathesis typology in Blevins & Garrett (2004), both of these latter types of metathesis occurring with CC clusters. Coarticulatory metathesis, which includes "the most common types of stop metathesis (PK > KP, TP > PT)," is "the result of extreme gestural overlap," which can occur in a number of different ways and cause excrescent segments, deletion or assimilation, and metathesis (Blevins & Garrett 2004, 126). The gestural overlap which creates metathesis involves "nearly simultaneous" "closure and/or release of two consonants with distinct articulatory gestures," with place cues being unrecoverable as a result (Blevins & Garrett 2004, 126). In such cases, C_1C_2 is reinterpreted as C_2C_1 "[i]f the righthand cluster edge contains unambiguous release cues"; these metatheses are unidirectional, turning labial-velar stop sequences into velar-labial stop sequences and coronal-noncoronal stop sequences into noncoronal-coronal stop sequences (Blevins & Garrett 2004, 126).

They note that in labial-velar stop sequences, the velar stop closure actually precedes the labial closure or occurs at the same time, though listeners will tend to perceive the velar closure before the labial closure regardless of when the closures happen articulatorily (Blevins & Garrett 2004). For Blevins & Garrett (2004), this explains the absence of KP > PK metathesis; regardless of the order, listeners will always perceive KP. Such is the case also with coronal-noncoronal stop metathesis: Closure for the noncoronal stop often precedes the closure for the coronal stop, and listeners, in any case, perceive the noncoronal before the coronal (Blevins & Garrett 2004).

This type of metathesis dovetails interestingly with coronal-noncoronal assimilation patterns, which are very common crosslinguistically, while the opposite patterns are not (Blevins & Garrett 2004). Blevins & Garrett (2004) argue that both phenomena occur because of the gestural overlap that occurs in coronal-noncoronal stop sequences, this being absent in noncoronal-coronal stop sequences and resulting in the absence of PT > TP metathesis.

Mokilese, a Micronesian language, provides one example of the unidirectionality of this metathesis. Every /pk/ cluster in the language can be "optionally realised as [kp]"; this happens with no other cluster, and /kp/ sequences are never realized as [pk] (Blevins & Garrett 2004, 136):

(27) /apkas/	[apkas], [akpas]	"now"
/kapkiːla/	[kapkiːla], [kakpiːla]	"to drop"
/dipkelkel/	[dipkelkel], [dikpelkel]	"to stumble"

Similarly, Aklanon, a Bisayan language spoken in the Philippines, has no [bg] clusters, these historical and underlying clusters having all become [gb] clusters; such did not occur, however, in Cebuano, another Bisayan language (Blevins & Garrett 2004, 136):

(28) Cebuano	Aklanon	
líbgus	lígbus	"mushroom"
palíbga	palígba (/pa-libug-a/)	"confuse him"

Blevins & Garrett (2004) also note the tendency of coarticulated labiovelars to change into velar-labial sequences, such as w > ɣʷ, w > gʷ (found in Indo-European languages); pʸ > k͡p, bʸ > g͡b, mʸ > ŋ͡m; pʸ > kʷ, bʸ > gʷ, mʸ > ŋʷ (found in Oceanic languages), which they argue is the result of the same phonetic factors: "the velar closure prior to labial closure as the jaw closes, and simultaneous or nearly simultaneous closure having the percept of velic closure" (137).

Similarly, in Cebuano, T{P,K} > P{K,T} occurs "as a regular sound change," with the opposite sound change unattested (Blevins & Garrett 2004, 138):

(29) Metathesis		No metathesis	
nm	inum : imn-a “drink”	mn	damán : damn-un “talk, walk in one’s sleep”
nŋ	tunúŋ : tuŋn-a “directly at a point”	ŋn	iŋún : iŋn-un “say, tell”
tp	atúp : atp-an, apt-an “roof”	pt	sáput : s-al-apt-un “bad temper”
tk	litik : litk-an, likt-an “snap the fingers”	kt	lakát : lakt-un “walk”
		kp	dakúp : dakp-an “arrest”

This is the equivalent coronal-noncoronal metathesis to the Mokilese labial-velar metathesis, and Blevins & Garrett (2004) point out that it is important that with /tp/ and /tk/ both [tp]~[pt] and [tk]~[kt] occur, because this suggests that “metathesis is directly related to degree of gestural overlap in the phonetic component” (138).

Auditory metathesis, on the other hand, occurs with “sibilant-stop or stop-sibilant sequences,” as in Modern Hebrew *hitpa’el* metathesis (Blevins & Garrett 2004, 127):

Table 3: Regular Sibilant Metatheses (Blevins & Garrett 2004, 138)

Language	Metathesis	Source
Old English	sk > ks	Weyhe 1908; Campbell 1959; Luick 1921-40; Jordan 1974
Faroese	sk > ks/__t	Lockwood 1955: 23-4
Lithuanian	coronal fricative + velar stop > k + fricative/__t	Seo and Hume 2001
Colloquial French	ks > sk/__#	Grammont 1923: 73
Savoyard	*ts > st/#__	Ultan 1978
Classical Aramaic languages	*t + sibilant > sibilant + t/V__V	Malone 1971, 1985, 1999
Ancient Greek	*dz > zd	Lejeune 1972: 113-116
Calabrian Greek	ps > sp	Rohlfs 1950: 74-6
Dutch	ps > sp	Stroop 1981-2

Blevins & Garrett (2004) theorize that “sibilant noise somehow distracts the listener, leading to high confusion rates with respect to the linear order of segments (Bregman 1990)” (128). This occurs because listeners segregate dissimilar speech sounds on different auditory streams, which, when combined, lead to indeterminacy of linear order (Blevins & Garrett 2004). They also suggest that “misperception of fricatives as affricates or stops, and vice versa” may also be a

"contributing factor" to this type of metathesis (Blevins & Garrett 2004, 128). The shorter the fricative is, the easier it becomes to misperceive it as a stop or affricate, which makes sense as all are obstruent sounds (Blevins & Garrett 2004). They also note that prosody may contribute to the directional symmetry in stop/fricative metathesis, with "longer sibilants…induc[ing] a greater confusion effect on segmental order and…thus more likely to undergo metathesis with an adjacent stop" (Blevins & Garrett 2004, 140).

Blevins & Garrett (2004) also note a number of important distinctions between their typological approach to metathesis and the phonetic optimization approach. Because their work is typological and phonetically-based, they predict gaps in the occurring patterns of metathesis, contrary to Hume (2004), who claims that any combination of sound could be a possible input or output of metathesis. One of these gaps is nasal-stop metathesis, which they argue is impossible because nasality and orality cannot "migrate across a neighbouring segment without directly affecting it; in such clusters assimilation is natural, but not metathesis" (Blevins & Garrett 2004, 140). Phonetic optimization, however, would predict that TN > NT metathesis should occur intervocalically, because "NT clusters are common (the only nongeminate clusters in some languages); and stop contrasts are relatively easy to perceive in prevocalic position" (Blevins & Garret 2004, 140). They argue that apparent metatheses like this are either not cases of metathesis at all or are due to other processes, like analogical morphophonology, this latter explanation accounting for "synchronic metathesis patterns" in East Cushitic and South Omotic languages (Blevins & Garrett 2004, 141).

Blevins & Garrett (2004) also point out that the phonetic optimization cannot account for the unidirectionality of PK > KP and T{P,K} > {P,K}T metatheses and predicts the unattested metatheses pg > gp and $V_1npV_2 > V_1nV_2p$ (141). For Blevins & Garrett (2004), directionality

patterns are the result of patterns of misperception, as "the various articulatory, acoustic, and perceptual factors underlying misperception and sound change are often intrinsically asymmetric" (Blevins & Garrett 2004, 142). In turn, their evolutionary phonology account predicts ɾV > Vɾ and Vɾ > ɾV metathesis, "a phonetically predictable transition…[being] reinterpreted as a full vowel, and [if] a historical vowel is reinterpreted as a transition," and auditory metathesis with other noisy sounds like /ɬ/ and clicks (Blevins & Garrett 2004, 141). These cases are unattested, but would provide further evidence for the evolutionary phonology explanation of metathesis. Importantly, Blevins & Garrett (2004) argue that types of metathesis which do not fit the attested typological patterns are the result of three possible "pathways other than sound change by which metathesis alternations may arise" (141). This includes loan adaptation, telescoping (with epenthesis and subsequent deletion), and analogical morphophonology, the morphologization of regular sound changes (Garrett & Blevins 2008).

One important final point from Blevins & Garrett (2004) relates to their overall evolutionary phonology theory. They "contend that reinterpretations of the ambiguities in real speech are the main force driving sound change" and that "the majority of attested regular historical metatheses in the world's languages can be explained as the result of phonetically natural sound changes" (Blevins & Garrett 2004, 143). For them, "phonetics determines emergent sound patterns" and "the typology of metathesis largely follows from convergent evolution, demonstrating the extent to which phonology is phonetically determined in the diachronic dimension" (Blevins & Garrett 2004, 143).

To conclude this section, the phonological study of metathesis has progressed from a rejection of the phenomenon as a sporadic change distinct from regular sound change, as the Neogrammarians believed; to formal analyses in generative grammar (either rule-based or OT-

based) which have tended to stress language phonotactics; and eventually to phonetically-based accounts in sophisticated models of language processing, perception, and change.

3.2 Research on Modern Hebrew Metathesis

Before I move on to a discussion of the experiment itself in §4, I would like to discuss two more papers written specifically about *hitpa'el* metathesis: Coetzee (1999) analyzes *hitpa'el* metathesis in Tiberian Hebrew within an OT framework similar to that proposed in Hume (1997) (and later echoed in Hume (2001)). Yanagawa (2003) analyzes the phenomenon in Modern Hebrew in OT, but she approaches the problem through articulatory phonology and the specifics of the articulation involved in producing stop + sibilant sequences.

Coetzee (1999) is a direct response to Malone (1993)'s analysis of *hitpa'el* metathesis, which is couched in a rule-based framework adapted from *The Sound Pattern of English*. Malone (1993) argues that *hitpa'el* metathesis can be accounted for with a rule he calls "Reflexive Fusion (RF)" (Coetzee 1999, 107):

(30) Reflexive Fusion (RF)

t	C		
[+ REF]	[+ HIT + ril - voi]		
1	2	→ 2	1

The distinctive features for 2 specify that only sibilants switch places with a preceding /t/ and only in *binyan hitpa'el*. Coetzee (1999) takes issue with this analysis in four ways: 1) The rule is not explanatory (i.e. there is no explanation of the metathesis takes place); 2) "Metathesis of stops and spirants seem [sic] to be a universal linguistic tendency" and Malone (1993)'s analysis does not take this into account, describing the process solely in terms of Tiberian Hebrew

grammar; 3) "In TH in general, a *t*-consonant never occurs in coda position followed by a syllable having a sibilant as onset," except in *binyan hitpa'el*, where Tiberian Hebrew morphology causes it to occur, so that the metathesis is not a special property of *binyan hitpa'el*, but rather a larger phonological process in the language; and 4) Coetzee (1999) argues that the metathesis should not be seen as the result of either a phonological or morphological process, but as "a consequence of a constraint on linguistic forms in TH" militating against t + sibilant clusters, both underlyingly (as evidenced by a lack of roots of the shape $C_1tC_{sibilant}$) and on the surface (107-109).

Coetzee (1999) postulates, based on Ultan (1978) and Hock (1985), that stop + sibilant sequences are marked, as they have a strong tendency to metathesize in many different languages. Ultan (1978) and Hock (1985), however, provide no theoretical explanation as to why stop + sibilant clusters should be marked. Coetzee (1999) notes that there are three possibilities as to why these clusters are marked: 1) Associational markedness, 2) lack of perceptibility of /t/ in coda position before a consonant, and 3) phonotactics. He argues against an approach to stop + sibilant markedness which invokes associational markedness and lack of perceptibility, because constraints against /t/ in coda position (*C/t) or against /t/ followed by a consonant (*tC) rule out existing un-metathesized *hitpa'el* forms (Coetzee 1999, 120-121). Reference must be made to the following sibilant, and Coetzee (1999) proposes that the constraint *t + SIB, which interacts with LINEAR, causes metathesis in Tiberian Hebrew. *t + SIB militates against t + sibilant sequences in order to avoid confusion between /ts/ and /t͡s/ sequences. Such an analysis suffers from the problem of the proliferation of parochial constraints that Blevins & Garrett (1998) lament; Coetzee (1999), in fact, points out that the form [notʃ-i] ("[t]he *qal* [*pa'al*] infinitive of the root n-t-š with a pronominal suffix added to it") is an exception to his analysis

and postulates there might be an intramorphemic LINEAR (and, indeed, this has been theorized as HOMLIN, homomorphemic linearity) or *t + SIB must apply only across morpheme boundaries (129). Taken to its logical conclusion, this means that there are two sets of constraints referring to every possible sound combination in any language: Those which apply within a morpheme and those which apply across morphemes.

Yanagawa (2003) argues for a gestural account of *hitpa'el* metathesis, noting that the metathesis only occurs word-medially and across a morpheme boundary. Using insights from articulatory phonology, which is concerned with articulatory gestures which interact with each other to produce phonemes, Yanagawa (2003) notes that, word-medially, much more gestural overlap occurs, leading to greater variability in gestural timing. Gestural timing is also more variable across morphemes, rather than within morphemes, and this instability allows for metathesis to occur (Yangawa 2003). Thus, word-initial and intramorphemic stop + sibilant sequences do not metathesize because they are more stable and cohesive than those in weaker positions (Yanagawa 2003). She also points out that this metathesis occurs with *hitpa'el* stop + sibilant sequences because sibilants do not overlap with following vowels, this also causing further gestural instability (Yanagawa 2003).

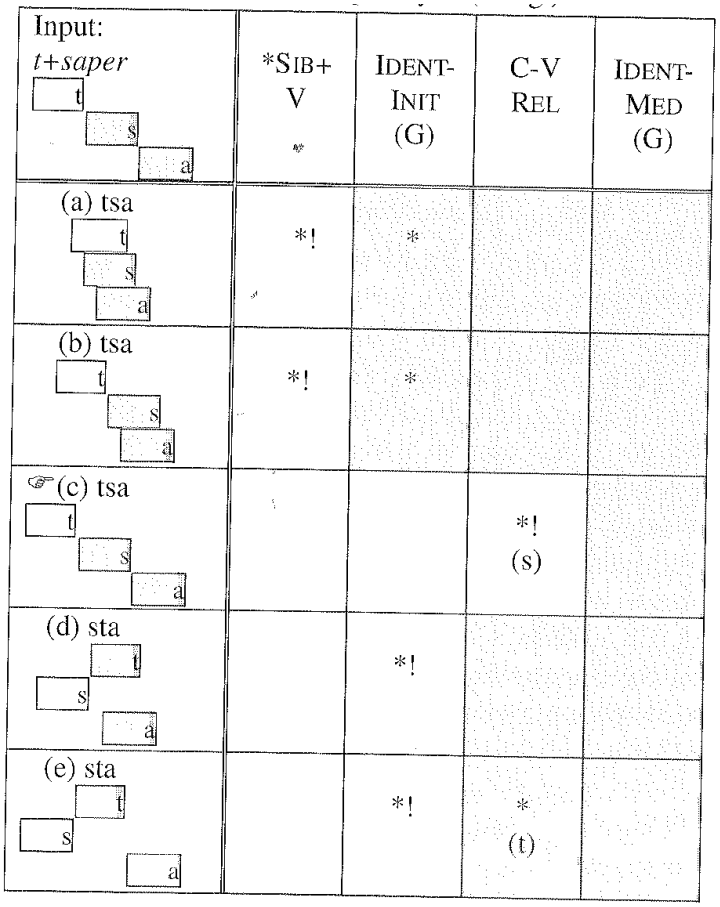
Yanagawa (2003) tested her hypothesis about the weaker environments by recording a female native speaker of Modern Hebrew reading 22 words: 6 of these were un-metathesized *hitpa'el* verbs with stop + stop and stop + fricative sequences; 8 were "*pi'el* verbs with word-initial hetero-morphemic target clusters"; and 8 were nouns and adjectives with mono-morphemic word-initial or word-medial consonant clusters (Yanagawa 2003, 1672). The *hitpa'el* verbs did show the greatest variability, because of their word-medial and hetero-morphemic environment (Yanagawa 2003). Because of this, Yanagawa (2003) proposes an OT account

which incorporates gestural overlap, arguing for the constraints C-V REL (C-V RELATION), "a vowel gesture must be overlapped with a consonant gesture in an onset," and *SIB+V (*SIBILANT + VOWEL), which "prohibits overlap between a sibilant and a vowel"; metathesis satisfies both constraints (Yanagawa 2003, 1672-1673). These markedness constraints are contrasted with two faithfulness constraints, IDENT-INIT(G) (IDENT-INITIAL (GESTURAL COORDINATION)) and IDENT-INIT(M) (IDENT-MEDIAL (GESTURAL COORDINATION)), which are violated by gestural overlap word-initially and word-medially. Together, these constraints form the ranking *SIB+V >> IDENT-INIT(G) >> C-V REL >> IDENT-MED(G) (Yanagawa 2003, 1673).

(31) Tableau 1: /hit+saper/ → [histaper] "he cut his hair" (Yanagawa 2003, 1673):

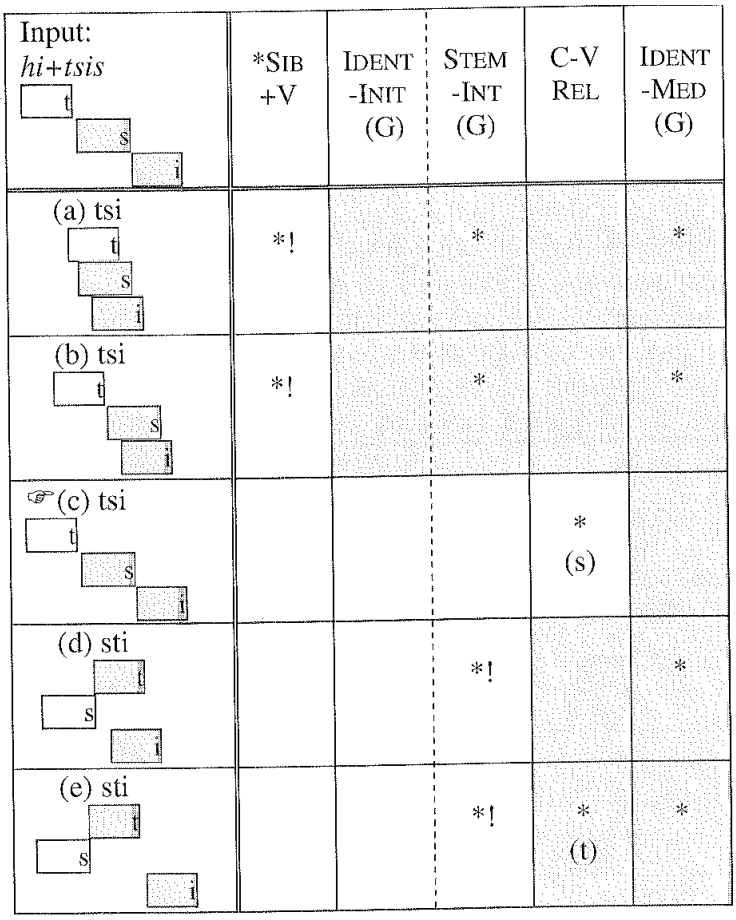
Input: *hit+saper*	*SIB+V	IDENT-INIT (G)	C-V REL	IDENT-MED (G)
(a) tsa	*!			*
(b) tsa	*!			*
(c) tsa			*! (s)	
☞(d) sta				*
(e) sta			*! (t)	*

(32) Tableau 2: /t+saper/ → [tsaper] "you (m.sg.) will tell" (Yanagawa 2003, 1673):



Finally, Yanagawa (2003) proposes STEM-INT(G) (STEM-INTERNAL COHESION (GESTURAL COORDINATION)) to account for the fact that metathesis does not occur intramorphemically. Together with the previous constraints, this forms the ranking *SIB+V >> IDENT-INIT(G), STEM-INT(G) >> C-V REL >> IDENT-MED(G).

(33) Tableau 3: /hi+tsis/ → [hitsis] "it fermented"



Yanagawa (2003)'s analysis is interesting because it suggests that there may be articulatory factors at play which may predispose stop + sibilant clusters to metathesis. However, I maintain, following Blevins & Garrett (2004), that sound change occurs because of misperception of ambiguous sound sequences, which are subsequently interpreted in non-historical ways. Another issue with Yanagawa (2003)'s analysis concerns whether a constraint like *SIB+V, which represents a phonetic reality (sibilants simply do not overlap very much with vowels), should even be represented in the phonology. Phonetic patterns certainly influence phonology, but should such low-level phonetic patterns actually be incorporated into the grammar?

4. Methods

In order to test the two main theories of metathesis, Hume (2004)'s indeterminacy/attestation model and Blevins & Garrett (2004)'s evolutionary phonology model, I created a speech perception experiment for native L1 English speakers using Modern Hebrew stimuli, these stimuli being appropriately degraded with English-speaking babble in order to both make the task more realistic and more difficult. It is important that I test speakers of English because this provides an empirical means of testing each theory on its fit with the data, since English and Modern Hebrew have different phonotactics/attested sound sequences. If English speakers can be shown to misperceive in the same way as Modern Hebrew speakers metathesize, this suggests that universal factors are at play in Modern Hebrew sibilant metathesis, a view in line with Blevins & Garrett (2004). On the other hand, if English speakers do not misperceive, or misperceive ambiguous sequences differently than Modern Hebrew speakers, this suggests that Hume (2004) is correct. In the following section, I describe in detail the methods for this speech perception experiment.

4.1 Predictions

Based on data from a pilot study (Jones 2015), which I refer to as Experiment 1, English speakers should misperceive stop-sibilant sequences as sibilant-stop sequences, resulting in the actual output of Modern Hebrew metathesis. In Jones (2015), I developed a Praat (Boersma & Weenink 2016) speech perception experiment which utilized eighteen non-word stimuli with characteristics similar to metathesized and un-metathesized Modern Hebrew *hitpa'el* verbs:

(34) List of experimental stimuli (Jones 2015):[13]

Sibilants: *[hitsa], [hista], *[hittsa], [hitsta], *[hidza], [hizda]
Stops: [hitka], *[hikta], [hitpa], *[hipta], [hitda], *[hidta]
Fricatives: [hitfa], *[hifta], [hitxa], *[hixta], *[hidva], *[hivda]

I specifically used non-word stimuli because I wanted English-speaking listeners, with no knowledge of Modern Hebrew, to focus specifically on the /tC/~/Ct/ alternation, to the exclusion of extraneous information (though this later proved to be unnecessary, as Experiment 2 demonstrates) (Jones 2015). The stimuli were controlled for voicing, so that there was only one voiced pair in each category, to the exclusion of the phonemes /l m n ʁ/; the stimuli were recorded by a female native speaker of Modern Hebrew in her mid-twenties using a Logitech noise-cancelling headset microphone (Jones 2015).

Once recorded, I combined the long sound file with a sample of multi-talker babble (from a sound file, "Crowd Noise," retrieved from http://www.mysoundeffect.com/humans/) using the Praat formula "self [col] + (2* Sound_soundb [col])" (Styler 2013, 38; Jones 2015). This has the effect of masking the stimuli in a 1:2 signal to noise ratio, making the task more difficult and more like a "normal" noisy speech environment, where speech perception would occur (Jones 2015). The recorded stimuli and the babble sample were recorded at different sampling rates, and it was necessary to sample them down to 22,500 Hz so they could be combined to create the final masked stimuli for Experiment 1 (Jones 2015). I then created individual stimuli from the long masked stimuli sound file by selecting 1.5 second snippets which included the individual stimulus plus babble noise at the onset and offset (Jones 2015).

Fifteen native English speakers, ranging in ages from 21 to 33, were tested in a multiple forced choice identification task (Jones 2015). Participants heard each of the eighteen stimuli in

[13] Non-existing combinations are marked with an asterisk.

four randomized blocks (so that each listener heard each of the stimuli four times, for a total of 72 X 15 = 1080 responses) at 10% of the total volume of the laptop; this had the effect of making the task even more difficult (Jones 2015).[14] The identification task itself consisted of responding to a stimulus with one of two possible choices: If a participant heard [hitka], for example, she would have to choose between the responses <hitka> on the left of the screen or <hikta> on the right, using the shift buttons on the keyboard to select which of the two choices she wished to answer with (Jones 2015).[15] If they wished to, participants could listen to a presented stimulus again by pressing the space bar, but they could do this only once (Jones 2015).[16] The first two blocks of Experiment 1 were separated from the last two by a rest period.

Though the overall confusion rate was low, with participants demonstrating between 2 and 15 individual misperceptions, results demonstrated relatively high levels of confusion for all groups of phonemes (Jones 2015). The most misperceived sequence was *[dz] → [zd] at a 100% confusion rate, which follows the pattern in Modern Hebrew, followed by [tp] → *[pt] (89.5%), and [tf] → *[ft] (80%) and [tx] → *[xt] (80%) (Jones 2015). *[ts] → [st] was also a highly misperceived sequence at 63.6%, this also following the Modern Hebrew metathesis pattern (Jones 2015). Confusion of *[tts] → [tst] was also higher than the opposite confusion, which follows the Modern Hebrew metathesis as well (Jones 2015).

[14] This experimental design likely introduces an implicit memory effect which detracts from the purpose of the experiment.
[15] There was no limit on how long a participant could wait and think about this decision, another flaw in the experimental design of this particular experiment.
[16] This is probably the most detrimental feature of the design of Experiment 1, because it allowed participants to listen to each stimulus twice, so that their initial perceptions were not recorded. This likely made the error rate much lower than it should have been.

Table 4: Confusion Rates (Jones 2015)

Direction of Confusion	Rate of Confusion	Number of Confusions
*hitsa → hista	63.6%	7/11
hista → *hitsa	36.4%	4/11
*hittsa → hitsta	53.8%	7/13
hitsta → *hittsa	46.2%	6/13
*hidza → hizda	100.0%	3/3
hizda → *hidza	0.0%	0/3
hitka → *hikta	57.9%	11/19
*hikta → hitka	42.1%	8/19
hitpa → *hipta	89.5%	17/19
*hipta → hitpa	10.5%	2/19
hitda → *hidta	40.0%	12/30
*hidta → hitda	60.0%	18/30
hitfa → *hifta	80.0%	4/5
*hifta → hitfa	20.0%	1/5
hitxa → *hixta	80.0%	4/5
*hixta → hitxa	20.0%	1/5
hidva → *hivda	50.0%	8/16
*hivda → hidva	50.0%	8/16

Indeed, when the direction of confusion for sibilants is compared to phonotactic probabilities for American English from Vitevitch & Luce (2004)'s online Phonotactic Probability Calculator, English speakers consistently misperceive in the same direction as Modern Hebrew speakers metathesize, even when this direction is *not* the direction expected based on English phonotactics.

Table 5: Phonotactic Probabilities for Stop/Sibilant Sequences (Jones 2015)

Phonotactic Sequence	Probability in English
t + s	5.5%
s + t	13.0%
t + t + s	15.1%
t + s + t	12.1%
d + z	5.3%
z + d	1.1%

The only exception to this is the sequence /st/, which is the more common sequence in English and which is the direction of metathesis/misperception in Modern Hebrew and in Experiment 1 (Jones 2015). However, if metathesis resulted in the more common attested patterns in English, [tst] → *[tts] and [zd] → *[dz] should be the directions of misperception for English speakers, but this is not the case. In Jones (2015), I claimed that these results were evidence against Hume (2004)'s indeterminacy/attestation model and that they provided indirect evidence for proposals like Blevins & Garrett (2004), which argue that sibilant metathesis is different from other metatheses and results from universal properties of the human auditory system (i.e., auditory stream decoupling), rather than language-specific mechanisms like attested phonotactic patterns.

Interestingly, the patterns of confusion for stops may also provide evidence for Blevins & Garrett (2004)'s theory. They theorize that T{P, K} > {P, K}T metathesis is unidirectional and that this metathesis only occurs in this direction because of the phonetics of the stop closures involved: In each case, the labial or velar stop closure is perceived by listeners as occurring prior to the coronal closure, resulting in metathesis of coronal and noncoronal (Blevins & Garrett 2004). The results of the Experiment 1 support this conclusion, as both [tp] and [tk] sequences were reanalyzed by participants as *[pt] and *[kt] (Jones 2015). These sequences do not metathesize in Modern Hebrew, but Blevins & Garrett (2004) provide good phonetic reasons why they should metathesize and why, in fact, they are predicted to metathesize in scenarios where ambiguity is high, like the context of Experiment 1.

Table 6: Phonotactic Probabilities for Stop/Stop Sequences (Jones 2015)

Phonotactic Sequence	Probability in English
t + k	7.2%
k + t	12.0%
t + p	6.5%
p + t	11.2%
t + d	5.3%
d + t	5.0%

Based on Experiment 1, which demonstrates that English speakers misperceive in the same direction as Modern Hebrew speakers metathesize, despite differing attested phonotactics, I predict that, in Experiment 2, English speakers should show the same stop/sibilant misperception patterns as Modern Hebrew sibilant metathesis. Ultimately, data like this would falsify Hume (2004)'s theory and lend support to an approach like Blevins & Garrett (2004), which ascribes sibilant metathesis to a universal, not language-specific, process, which operates the same way in every language, regardless of attested patterns in the language.

4.2 Stimuli

The stimuli for Experiment 2 consist of 198 verbs in *binyan hitpa'el* drawn from Bolozky (2008)'s *501 Hebrew Verbs*. There are ten metathesized and ten un-metathesized verbs for each of 10 phoneme categories (except for /l/, which only has nine verbs each, as explained below) for a total of 198, including both existing and non-existing forms. Each of the verbs are grouped by the initial consonant in the verbal stem, according to the following schema:

(35) Stops: /p/, /k/
Non-sibilant Fricatives: /ʁ/, /x/
Sibilant Fricatives: /s/, /z/, /t͡s/
Sonorants: /l/, /m/, /n/

Grouping the verbs in this way allows for comparison across different manners of articulation, so that any tendencies to misperceive in particular directions can be compared with tendencies in different phonemes. Initially, I also included /d/ with the stops, but these items were ultimately thrown out because my speaker had trouble articulating them (either consistently switching linear order or pronouncing a /td/ sequence as [dd]). The stimuli for Experiment 2 differ from the stimuli for Experiment 1 because they represent existing Modern Hebrew verbs, with their metathesized or un-metathesized (for sibilants) counterparts, and, thus, they represent most of the phonemes in the language, with only those verbal stems which cannot alternate excluded (/h/ and /t/, for example, because *[hiht…] is probably not recoverable and /hitt…/ assimilates to [hit…]).

Because I am most interested in the sibilant fricatives, I have attempted to balance each of the other categories based on the sibilants. There are four sibilant fricatives in Modern Hebrew, /s z ʃ t͡s/, but /ʃ/ is excluded from Experiment 2 because it might be interpreted by English speakers as the phoneme /t͡ʃ/ when it follows /t/ in the non-existing unmetathesized *hitpa'el* verbs used. This would ultimately be detrimental to the experiment, because I would have no way of determining what English-speaking listeners are hearing when they hear /t/ + /ʃ/ sequences. Because the included sibilant fricatives /s/, /z/, and /t͡s/ include a voiced consonant, each of the non-sonorant phoneme categories include a voiced consonant, in order to control for voicing. Because of the constraints of Modern Hebrew, there are simply not enough non-sibilant fricatives to include more than two phonemes in this category. I tried to choose different places of articulation as much as I could, in order to vary the stimuli, but, because I am working with natural language data, this could not always be done. Any imbalances are simply due to the individual phonetic, phonological, and phonotactic characteristics of Modern Hebrew.

I attempted to choose verbs from the most common group of *hitpa'el* verbs, which consists of a triconsonantal stem with the vocalic pattern {ae}, represented by the citation form [hitpael] (with the second consonant in the stem of this citation form traditionally being /ʕ/, which is no longer pronounced, except as /ʔ/ in very careful, formal speech). In order to have enough verbs in each category, I included some verbs which come from Mishnaic or Medieval Hebrew and are probably no longer used, except in old texts. A few verbs are really *nitpa'el* verbs, a verbal conjugation which resulted from confusion with the *nif'al* conjugation, which is generally passive and consists of a prefix /n-/, which replaced the initial /h-/ of the /hit-/ prefix of *binyan hitpa'el* (Berman 1978). These forms eventually died out, but *nitpa'el* verbs are conjugated in exactly the same way as *hitpa'el* verbs in the past tense, so I included a few of them in the stimuli, where it was necessary to achieve the desired number of verbs in each category.

Overall, there were enough verbs in each category except for /d/, /z/, and /l/, which simply did not have enough forms. For these phonemes, I had to include possible, but non-existing forms which I created, and different phonological instantiations of *binyan hitpa'el* including verbs with less than three consonants in the stem; quadriliteral stems; and stems with different vocalic patterns, like {aa}. Bolozky (2008), for example, only lists seven *hitpa'el* verbs with /d/ as initial stem consonant. In order to flesh out the category with more verbs, I took /d/-initial stems which exist in other *binyanim*, like [daxaf] "push," and fitted them into *binyan hitpa'el*, to come up with possible, but non-existing verbs like [hitdaxef], which might mean something like "push oneself (to do something)" or "push each other." Similarly, I included the verb [hitdaeg] "become worried/afraid," which lacks a second stem consonant because the orthographic /ʔ/ is no longer pronounced (Bolozky 2008, 117). Coincidentally, [hitdaeg] is also a

nitpa'el verb (Bolozky 2008). As previously mentioned, those items with /d/ were thrown out because the speaker could not pronounce them consistently, as well as the fact that most of them were non-existent verbs.

I also initially included some quadriliteral stems, like [hizdalzel] "become a laughing-stock; make oneself cheap," which do not fit the general three consonant pattern of *binyan hitpa'el*, but which nevertheless still use the typical {ae} vocalic pattern (Bolozky 2008, 159). I excluded this item later, because I actually had 11 /z/ pairs. I also threw out one quadriliteral /l/ pair, [hitlaxlex]~*[hiltaxlex], due to the fact that my speaker could not pronounce the metathesized version (producing something like *[hiltxalex] instead). Finally, I included verbs like [hitlakax] "catch fire; flare up (quarrel)," which, while triconsonantal, have a different vocalic pattern, like {aa}, than most *hitpa'el* verbs usually have (Bolozky 2008, 349). The {aa} pattern in [hitlakax] is the result of the final guttural consonant, which, historically, was a pharyngeal /ħ/ that caused an expected final /e/ vowel to lower to /a/. My speaker did not, in fact, pronounce any of these items with different vowel patterns, so that the total number of items did not actually include any deviant vowel patterns. This was an unexpected but welcome result, as it evened out all of the stimulus items. My speaker may have treated the stimulus items as if they were non-words (even for those that are real words in Modern Hebrew, possibly because of the artificiality of the reading task), substituting the default {ae} vocalic pattern. This outcome could also simply be the result of analogical levelling: Some *hitpa'el* verbs, mostly those with guttural consonants, have alternative forms with the more common {ae} vowels, as Bolozky (2008) notes frequently. Out of a total of 198 verbs, then, all are triconsonantal and share the same {ae} vocalic pattern.

I recorded a male native speaker of Modern Hebrew saying the 198 stimuli from a list, recording in a quiet room in the School of Middle Eastern and North African Studies and in his quiet office, using a Logitech noise-cancelling headset microphone and Praat (Boersma & Weenink 2016), a phonetics program for recording, reviewing, and manipulating speech data. The stimuli were recorded in two separate sessions, in order to correct pronunciation mistakes (metathesized forms, instead of un-metathesized forms, for example). Once the stimuli were recorded, I textgridded the long sound files in Praat, labeling each of the stimulus items, along with any undesired speech noises (lip smacking, throat swallowing, and stuttering, for example) and long silent periods. After textgridding, I isolated each individual stimulus and extracted it from the large sound file using a Praat script. Each individual stimulus sound file was around the same length, although I later accounted for deviations by introducing the same amount of silence to the beginning portion of each stimulus file with one of Matthew Winn's Praat scripts (available at http://www.mattwinn.com/praat.html), which normalized duration. In order to make sure each of the items also had similar amplitudes, so the stimulus files were all roughly the same loudness, I also normalized amplitude with another of Matthew Winn's Praat scripts.

In order to make the identification task more difficult, and in order to replicate the noisy conditions which occur every day in real life, I combined the stimuli files with a sample of a speaker in multi-talker babble (Thibodeau n.d.). This speech file consists of a male speaker of American English saying the sentence, "I can hurl tennis rackets at small moving objects with deadly accuracy," (from comedian Hugh Gallagher's humorous college application essay) against a background of multi-talker babble at 60 dB, into a Comtek FM environmental microphone on his body while he is twelve feet away from a hypothetical listener (Thibodeau n.d.). In order to determine the best combination of stimulus and noise, I tested different signal-

to-noise ratios (SNR). I began with a difference of 0 dB (a 1:1 SNR) and created test stimuli with different decibel levels both above and below the initial reference point (+1 dB and -1 dB, for example). Ultimately, I decided on an SNR of -6 dB; this 1:2 signal to noise ratio was meant to overwhelm the acoustic characteristics of the stimuli, which I verified by viewing the spectrogram in Praat. The spectral characteristics of the stimuli were obscured by the babble, making it difficult to visually see where the stimuli were within the noise. This effectively means that the human auditory system should have trouble "seeing" the spectral characteristics of the original stimuli as well.

4.3 Participants

Twenty-one monolingual English speakers participated in Experiment 2. The monolingual English speakers were drawn from the Department of Linguistics undergraduate subject pool at the University of Arizona, as well as undergraduate students who responded to recruitment emails sent out on the listservs of several of the constituent departments of the College of Social and Behavioral Sciences and the Eller College of Management. The linguistics subject pool consists of undergraduate students at the University of Arizona who are currently enrolled in linguistics classes and who are either required to participate in experiments in the linguistics department as part of their final grade or have the ability do so for extra credit. To be included in the study, participants had to be monolingual speakers of English who did not speak, and had never studied, a Semitic language (either Modern Hebrew or Arabic, for the purposes of this study). Each participant drawn from the linguistics subject pool was compensated for their participation in the study with extra credit. The undergraduates drawn from other departments

were compensated with $5 gift cards to Caffe Lucé, a popular café on campus, for their participation in the study.

Ideally, I wanted to have an equal number of participants who spoke English and Modern Hebrew, in order to investigate how Modern Hebrew speakers would perceive the experimental stimuli, but this was simply not possible. There are few speakers of Modern Hebrew in Tucson, Arizona, and I was limited by this fact in the number of Hebrew-speaking participants I could recruit into my experiment. As I mentioned previously in §1, there are a number of other issues with testing Modern Hebrew speakers, which would have necessitated changing the experiment, and these important complications prevented me from testing Modern Hebrew speakers in the same experiment as English speakers.

4.4 Procedure

I created a forced choice identification task using DMDX[17] (Forster & Forster 2003), an experimental psycholinguistic program created by Prof. Kenneth Forster and Jonathan Forster of the Department of Psychology at the University of Arizona (www.u.arizona.edu/~kforster/dmdx/dmdx.htm). This user-friendly experimental software is used by psychologists and linguists all over the world to perform psycholinguistic experiments and easily record millisecond accurate reaction times. DMDX's AutoMode removes the need to set up the program using a special program called TimeDX, which tests that the computer has the requisite requirements and components to run a reaction time experiment, ensuring that the experiment runs the same on different devices. Experimental instructions are provided by a Rich

[17] The name "DMDX" is an abbreviation of "**DM**ASTR **D**irect**X**," which is derived from DMASTR, the family of DOS programs that DMDX is a member of, and DirectX, a software program component which allows for DMDX's excellent timing accuracy.

Text Format (.rtf) input file, which instructs the program what to present and when to collect reaction times.

Experiment 2 consists of four scrambled blocks of stimuli, each block consisting of an equal number of stimulus items from each phoneme category (5 items X 10 phoneme categories = a total of 50 items per scrambling block), except for the last block, which only includes three items from the /l/ category (for a total of 48 items), because the last pair ([hitlaxlex]/[hiltaxlex]) was thrown out. This pre-scrambling of the items in each scrambling block, which are scrambled amongst themselves during the actual experiment, ensures that there are an equal number of all stimulus items in each block, so that participants do not hear a disproportional number of stimuli from any one phoneme category during the experiment. Stimulus items in the .rtf file are coded to specify each of the 198 stimulus items individually and for what kind of response is anticipated (+ or -). [hitpageʃ], for example, is coded +11011001; the "+" represents that a left shift response is correct ("-" representing a right shift response), while the first number in the sequence represents whether the stimulus item is unmetathesized (1) or metathesized (2). The second number represents the phoneme category (1 = stop, 2 = non-sibilant fricative, 3 = sibilant fricative, 4 = sonorant). The third and fourth numbers represent the individual phoneme (01 = /p/, 02 = /k/, 03 = /ʁ/, 04 = /x/, 05 = /s/, 06 = /z/, 07 = /t͡s/, 08 = /l/, 09 = /m/, 10 = /n/), while the fifth number represents voicing (1 = voiceless, 2 = voiced). The sixth, seventh, and eighth numbers represent the individual item number (001-198).

Before beginning the Experiment 2, participants were consented with the Douglass Phonetics Laboratory's consent form and then verbally told the experimental directions, which were also written at the beginning of the experiment for participants to review. This included a description of the task and the participant's role in the experiment, namely that the experiment

was a listening task, that participants would be required to respond as quickly as possible (guessing if necessary) by using the left and right shift buttons, that the responses were words written in English orthography, and that the experiment was broken up into four blocks with rest periods in between each block, during which participants could stop to rest before moving on. Participants heard a stimulus and were required to respond with one of two response choices (with the left or right shift buttons) in less than 2.5 seconds, at which time the experiment timed out and moved on to the next stimulus item. For example, if a participant heard the stimulus item [hitkabel], they had to respond with either <hitkabel> or <hiktabel> within 2.5 seconds, otherwise the program moved on and an incorrect response of -2500 was recorded. They continued to respond until the end of one of the 50 item blocks, at which point they could take a break before moving on to the next block. After the four blocks were completed, either I or the participant pressed escape and saved the data, which was automatically added to the bottom of an .azk file, which recorded the response times and whether the response was correct or incorrect, along with a running tally of how long the experiment took and when each response was recorded during the time of the experiment. The total experiment took between 35 and 40 minutes, with five to ten minutes for directions. Participants were scheduled one after another in one hour blocks, starting on the whole hour.

4.5 Analysis

I chose to use DMDX, rather than Praat, because DMDX collects millisecond accurate reaction times, which Praat simply cannot, these reaction times being continuous data that can be easily statistically analyzed. The data from Experiment 2 were analyzed in R using logistic regression, as well as a binomial linear mixed effects regression model; the results of each of

these statistical analyses were then compared to each other in order to interpret the results. Phonotactic frequencies for English and Modern Hebrew were compared to each other, using the MRC Psycholinguistic Database (Wilson 1988) for English and a digitized version of *Milon Even-Shoshan* (Even-Shoshan Dictionary, Even-Shoshan 2003) for Modern Hebrew, in order to determine whether English speakers consistently misperceive ambiguous *hitpa'el* verbs containing sibilants in ways unexpected from the attested patterns in English. Hume (2004) uses the MRC Psycholinguistic Database (Wilson 1988) to argue for her assertion that metathesis always creates a phonotactic combination that is the more common combination in a language. I have chosen to use the MRC Psycholinguistic Database (Wilson 1988) as well, for the purpose of comparison to her method of determining the more commonly attested sequences in a language, and because the MRC Psycholinguistic Database (Wilson 1988) has an easy-to-use interface which supplies the relevant information that I need for this experiment. The MRC Psycholinguistic Database (Wilson 1988), housed at the University of Western Australia, allows users to search for specific sequences of sounds in different combinations, among other search methods. In order to determine the phonotactic frequencies of each of the relevant combinations of sounds in my study, I searched in the database for all words which had the combinations word-initially, word-medially, and word-finally. This method works well, although, for those sound combinations which occur very frequently (at least two of which are included in this study), the database truncates the list at 5000 words. Determining the number of words in the database with the more frequent combinations, then, is hampered by the fact that the database search engine does not display every relevant word, but only the first 5000. In the following section (§5), I will discuss the statistical tests used to interpret the results, as well as compare the results to the phonotactic frequencies determined for English and Modern Hebrew.

5. Results

Of the twenty-one participants tested in Experiment 2, I excluded three of the participants from the final analysis, leaving a total of eighteen participants included in the results. One of the three excluded participants had a diagnosed hearing problem, and another participant adjusted the volume on the computer from the required 20% of total volume to 34%, despite explicit instructions to the contrary. The third excluded participant listened at twice the required volume and had a 50% error rate, performing exactly at chance during the experimental task, and was excluded on this basis. Interestingly, both the participant with a diagnosed hearing problem and the participant who listened to the experiment at 34% volume fell within the error rate range of the eighteen included participants, the former achieving a confusion rate of 45% (one percent lower than the worst participant, at 46%) and the latter achieving a confusion rate of 13% (one percent higher than the best participant, at 12%). The overall error rate for all participants was 27.06%, a more satisfying result compared with Experiment 1. Individual error rates ranged from 12% to 46%. Those participants who had some language experience appeared to do better at the task than those without (for example, Participant 13, a fluent L2 speaker of German, achieved a 14% error rate), though this was not borne out by every participant (Participant 22, a fluent L2 speaker of Spanish, achieved a 32% error rate, greater than twice the error rate of Participant 13), and I did not investigate any further, since the influence of language ability on confusion rates is not a central question of this experiment.

Table 7: Participant Results

Participant	Sex	Age	Error Rate
03	M	22	20%
05	F	20	24%
06	F	33	43%
07	M	22	34%
09	F	22	21%
10	F	19	46%
11	F	32	12%
12	F	24	26%
13	F	22	14%
14	M	19	23%
15	F	24	37%
16	F	19	19%
17	F	27	29%
19	M	21	15%
20	F	19	40%
21	M	25	18%
22	M	31	32%
23	M	33	34%

Using RStudio (RStudio Team 2015), a statistical software interface for the widely-used programming language R, I performed two exploratory analyses of the data. First, I performed a logistic regression, and, secondly, I performed a linear mixed effects regression, in order to fully investigate the data. The logistic regression, using reaction time (Resp) as the dependent variable and metathesis (Meth), phoneme type (Seg), and voicing (Voice) as independent variables, found that only phoneme type is significant, with voicing barely significant. Fricative, sibilant, sonorant, and stop phonemes are significant ($p < 0.0001$) (with fricative phoneme type as the baseline), for determining the type of response which participants gave, while voicing is less significant ($p < 0.05$). Metathesis is not significant; this is to be expected since I am claiming that metathesis is only significant for sibilants, misperception occurring with these segments disproportionately due to their acoustic characteristics, which cause listeners to misinterpret linear order to a larger extent. The significance of the segments involved is important, because

this includes sibilants, potentially pointing to the importance of the type of segment in causing misperception.

In order to further refine my analysis of the results, I also performed a linear mixed effects regression (LMER) in R. LMER models also include random effects, in addition to the fixed effects included in logistic regression models, so they can potentially capture important effects which might otherwise be missed by simpler models. This attests to their increasing popularity in linguistics and many other scientific fields. The LMER for Experiment 2 determined, in line with the logistic regression, that metathesis is not significant, but that phoneme type is. The LMER and logistic regressions differed, however, on whether voicing is significant, with the LMER determining that voicing is significant.

Based on the breakdown of the LMER, it is important to note that fricatives, sibilants, and stops are the most significant categories within the phoneme type category ($p < 0.0001$), while sonorants are much less significant than the former phoneme types ($p < 0.05$). This is pertinent to my hypothesis because the LMER demonstrates that sibilants are significantly involved in the misperception, while other phoneme types are not. The LMER also reveals that voicing is highly significant ($p < 0.0001$), meaning that voicing also determines the response that participants give. Interestingly, metathesis only becomes significant ($p < 0.01$), if it is analyzed together with voicing in a separate LMER.

Reaction time data is also illuminating. Plotting the reaction times of correct responses reveals that participants took the longest time to identify fricatives correctly, followed by sibilants, sonorants, and stops. Participants took much longer identifying fricatives and sibilants than they did sonorants and stops, the latter being much closer in reaction times, suggesting that fricatives (including sibilant fricatives) are much more perceptually difficult for listeners than

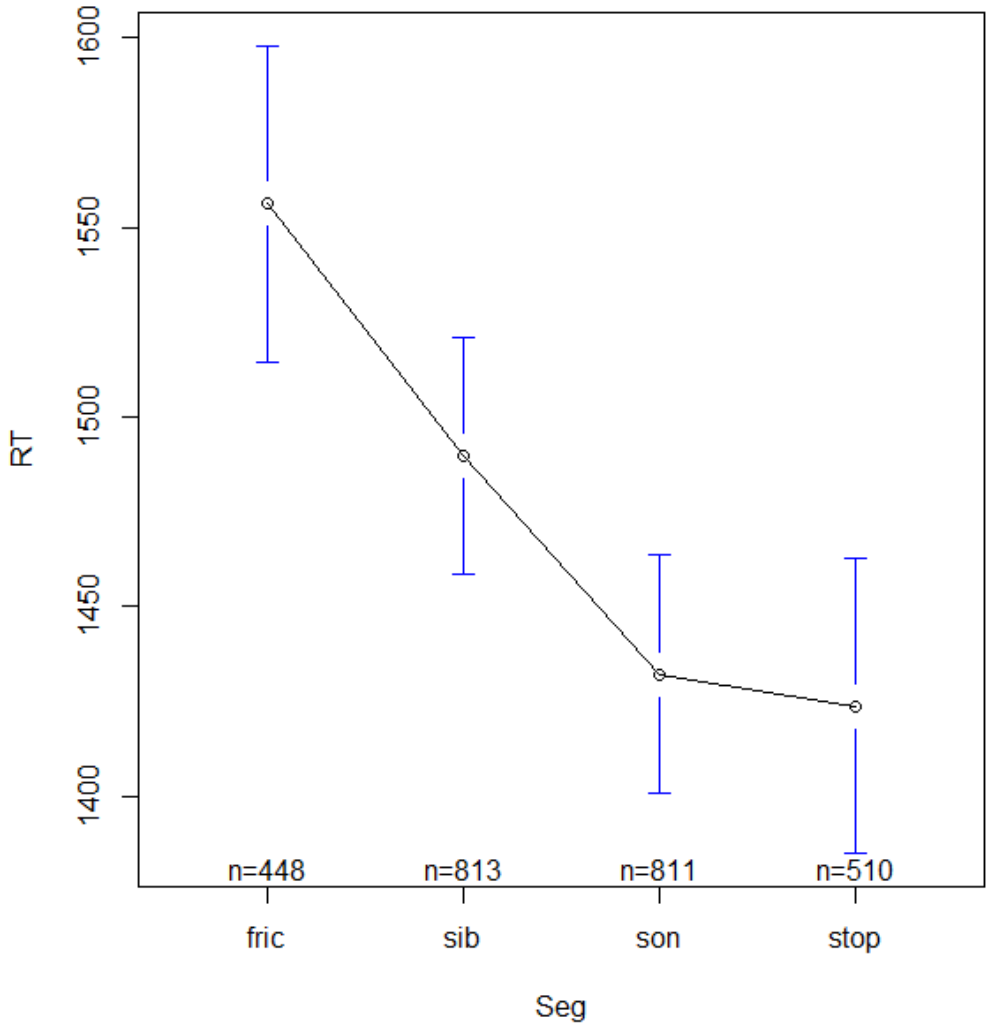


Figure 1: Reaction Times for Correct Responses Across Phoneme Type

either sonorants or stops, perhaps explaining why fricatives and sibilants are so often involved in metathesis in the first place. Plotting reaction times for voicing reveals additional information, demonstrating that voiceless sounds are more difficult to perceive correctly (i.e. reaction times for voiceless sounds are longer) than voiced sounds. Voiceless fricatives and sibilants, then, are

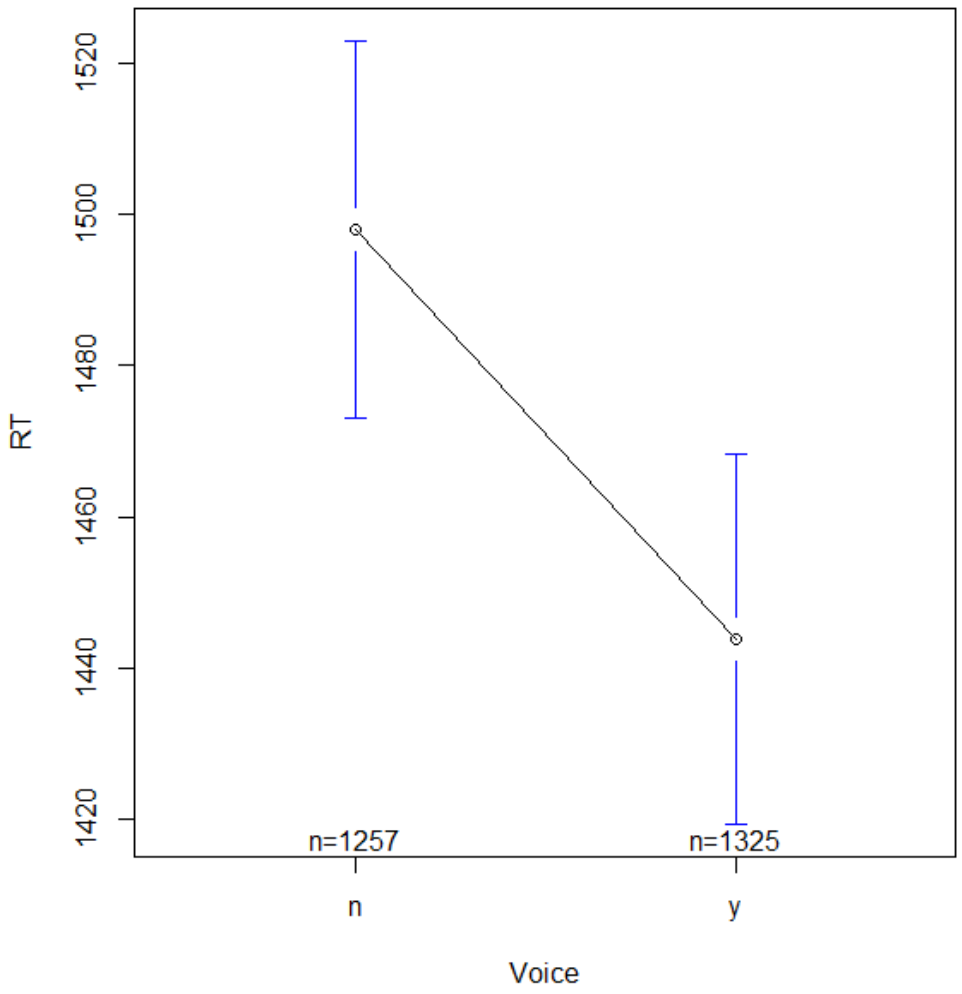


Figure 2: Reaction Times for Correct Responses Across Voicing

the most difficult sounds for participants to perceive.

When the reaction times for correct responses are plotted by phoneme type and metathesis, the most interesting and relevant information is revealed. Both stops and sonorants

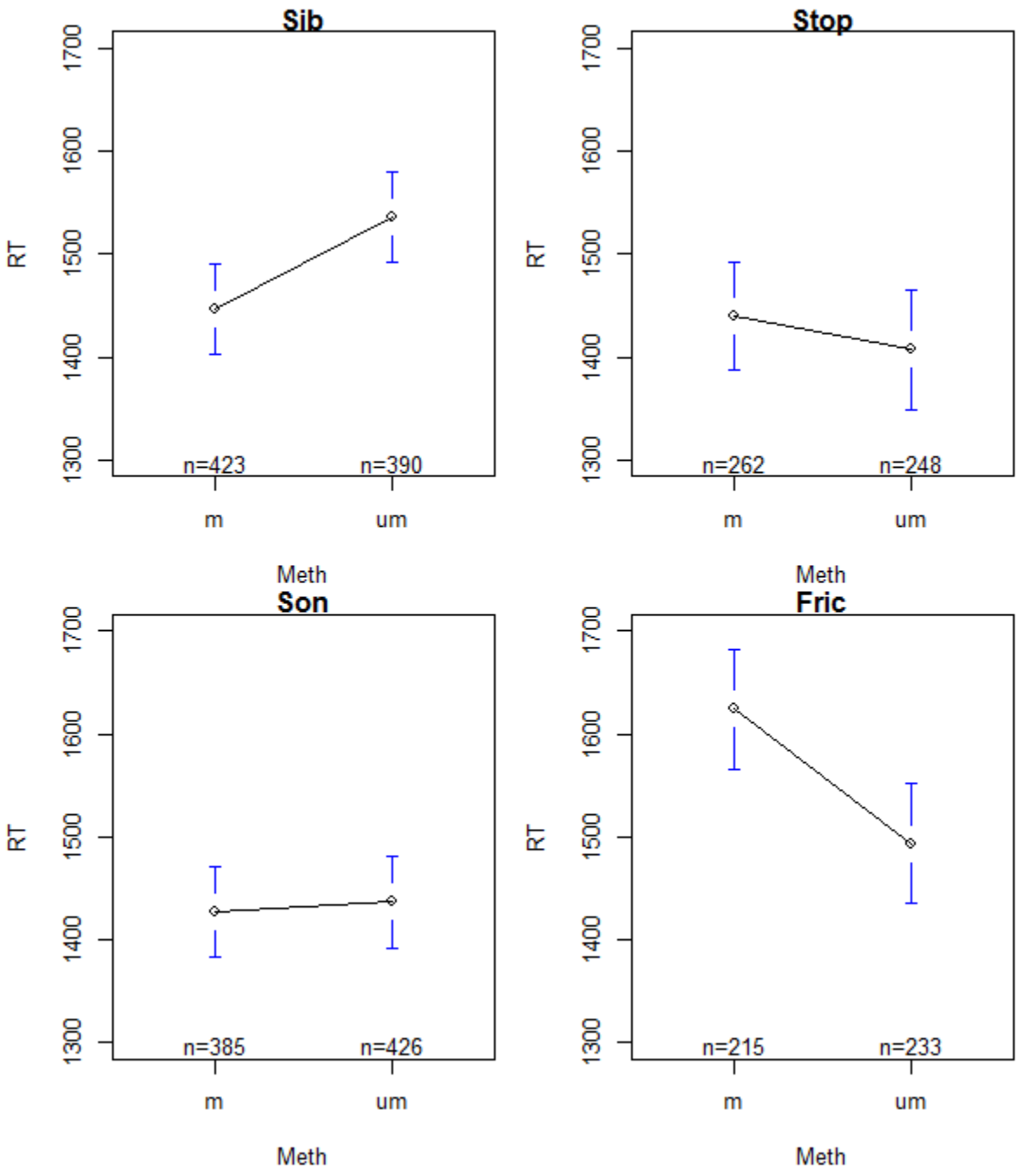


Figure 3: Reaction Times by Phoneme Type and Metathesis

have almost no slope, which means that there is little difference in reaction times for the metathesized and un-metathesized stimuli. The fricatives have a negative slope, as participants had more difficulty identifying metathesized stimuli than they did un-metathesized stimuli. However, the sibilants have a positive slope; only with the sibilants did participants have more

difficulty identifying the un-metathesized stimuli than the metathesized stimuli. In other words, sibilants are uniquely misperceived, being the only phoneme type that participants took longer to respond correctly to the unmetathesized stimuli, and constitute a separate and distinct group from the fricatives, as far as misperception is concerned, as Blevins & Garrett (2004) point out in their typology of metathesis.

5.1 Direction of Misperception and Phonotactic Frequency

In the case of the sibilants, English speakers misperceived the un-metathesized sibilant sequences as metathesized sibilant sequences in the majority of cases. This is the same direction of misperception as the metathesis that occurs in Modern Hebrew.

Table 8: Confusion Rates

Direction of Confusion	Rate of Confusion	Number of Confusions
tp → *pt	**62.5%**	65/104
*pt → tp	37.5%	39/104
tk → *kt	44.3%	47/106
*kt → tk	**55.7%**	59/106
tʁ → *ʁt	29.8%	37/124
*ʁt → tʁ	**70.2%**	87/124
tx → *xt	**57.4%**	89/155
*xt → tx	42.6%	66/155
*ts → st	**53.4%**	31/58
st → *ts	46.6%	27/58
*dz → zd	**57.3%**	47/82
zd → *dz	42.7%	35/82
*t͡ts → t͡st	**57.1%**	72/126
t͡st → *t͡ts	42.9%	54/126
tl → *lt	26.6%	29/109
*lt → tl	**73.4%**	80/109
tm → *mt	47.7%	31/65
*mt → tm	**52.3%**	34/65
tn → *nt	**62.1%**	36/58
*nt → tn	37.9%	22/58

The results for /z/ and /t͡s/ are both 57%, while the results for /s/ are 53%, which is slightly lower and closer to 50%. I would argue that this is due to the presence in English of both /st/ and /ts/, with /ts/ serving as a good competitor, causing a greater percentage of /st/ sequences to be misperceived as /ts/. Indeed, reviewing the results from the MRC Psycholinguistic Database (Wilson 1988) appear to confirm this.

Using the MRC Psycholinguistic Database (Wilson 1988), I calculated phonotactic frequencies for each of the phoneme combinations that I tested in the experiment, with the exception of /tx/~/xt/ and /ʁt/~/tʁ/, because these sounds do not exist in English. These are listed below in Table 9.

Table 9: Phonotactic Frequencies for English

Sound Sequence	Phonotactic Frequency	Rate of Confusion
tp	5.6%	37.5%
pt	**94.4%**	**62.5%**
tk	2.9%	**55.7%**
kt	**97.1%**	44.3%
ts	35.3%	46.6%
st	**64.7%**[18]	**53.4%**
dz	**72.5%**	42.7%
zd	27.5%	**57.3%**
t͡ts	0%	42.9%
t͡st	**100%**	**57.1%**
tl	**60.0%**	**73.4%**
lt	40.0%	26.6%
tm	**93.6%**	**52.3%**
mt	6.4%	47.7%
tn	7.8%	37.9%
nt	**92.2%**[19]	**62.1%**

The phonotactic frequencies are revealing because they do not adequately predict English speakers' reactions to ambiguous stimuli, as would be expected from Hume (2004)'s model.

[18] The database cut off the list of returned words meeting the criteria at 5000.
[19] Ibid.

Particularly, the phonotactic frequencies reveal that /dz/ is the far more common sound sequence in English, at a nearly 3:1 ratio, despite the fact English speakers misperceive /dz/ as /zd/ 57.32% of the time. This is directly contrary to what the indeterminacy/attestation model would predict.

In explaining the direction of metathesis in Modern Hebrew, Hume (2004) argues that the /St/ sequence is more common in the language, explaining speakers' tendency to metathesize in that direction. This is borne out by my analysis of *Milon Even-Shoshan* (Even-Shoshan 2003), which is displayed below in Table 10.

Table 10: Phonotactic Frequencies for Modern Hebrew (Even-Shoshan 2003)

Sound Sequence	**Phonotactic Frequency**
ts	1.6%
st	98.4%
dz	3.3%
zd	96.7%
t͡ts	4.0%
t͡st	96.0%
tʃ	5.4%
ʃt	94.6%

At first glance, Hume (2004)'s argument makes sense, but what it fails to take into account is that the huge percentages of /St/ sound sequences already include *binyan hitpa'el*. The main reason that /St/ sequences are so common is because of the prevalence of metathesized *hitpa'el* verbs. In other words, it is a circular argument to claim that sibilant metathesis occurs in Modern Hebrew because speakers are metathesizing ambiguous sound sequences into the more common sound sequence; the more common sound sequence is only the more common sequence because of metathesis. Some other phenomenon must cause this metathesis.

6. Discussion

The results of Experiment 2 reveal two important findings: 1) Stop-sibilant sequences are uniquely misperceived by English speakers as sibilant-stop sequences, a pattern distinct from all other phoneme types investigated and which is superficially similar to the direction of metathesis in Modern Hebrew; and 2) this misperception cannot be explained solely by reference to the most common phonotactics of English, with speakers misperceiving /dz/ as /zd/ in the majority of cases, despite /dz/ being the more common sequences in English. Ultimately, these findings provide support for Blevins & Garrett (2004)'s theory of metathesis and falsify Hume (2004)'s indeterminacy/attestation model.

6.1 Contra Hume (2004)

Hume (2004)'s model of metathesis involves two main motivating factors for metathesis: Indeterminacy in the speech signal and attestation of sequences in a language. Indeterminate sequences of sounds, which listeners have trouble deciphering, are converted into sequences of sounds which are common in a particular language; in other words, listeners have knowledge of the phonotactics of their own languages and the more common sequences of sounds and bring this knowledge to bear when interpreting speech, which may be indeterminate by virtue of similar sounds with masked phonetic cues or by virtue of stretched out phonetic features which lead to overlap and reinterpretation of the indeterminate sequence. Her theory goes a long way in explaining why metathesis occurs, but it does not adequately capture the facts of this experiment or all the facts of metathesis in general, as Blevins & Garrett (2004) point out.

Experiment 2 demonstrates that an ambiguous (or indeterminate, in Hume (2004)'s terminology) speech signal is not always reinterpreted as the more common sequence in a

language. The results from the MRC Psycholinguistic Database (Wilson 1988), which Hume (2004) also used to determine the frequencies of sound sequences in her paper, demonstrate that /dz/ is the more common sound sequence in English, at 72.5%, with only 27.5% of words having the less common sequence /zd/. However, the perception results from Experiment 2 reveal, despite this attested pattern in English, that English speakers misperceive /dz/ as /zd/ at a higher rate (57.3%) than they misperceive /zd/ as /dz/ (42.7%). These results are directly contrary to Hume (2004)'s model and cannot be adequately explained. Hume (2004) explicitly states that when two sequences are attested in the language, the listener will choose to reinterpret an indeterminate sequence into the "most robust sequence, that is, the one with the highest frequency," although this did not occur with /dz/~/zd/ sequences (222). English-speaking listeners were biased toward the less common sequence in their language, rather than reinterpreting the sequence into the more common, more robust sequence. Similarly, this also occurred with the /tk/~/kt/ alternation, where English speakers misperceived the more common sequence (/kt/) as the less common sequence (/tk/). This case contradicts Hume (2004)'s model even more, because the sequence /tk/ only occurs in 2.9% of English words, while /kt/ occurs in 97.1% of English words, by far the more common sequence. Yet English speakers misperceived /kt/ as /tk/ 55.66% of the time, while they only misperceived /tk/ as /kt/ 44.34% of the time. As with the /dz/~/zd/ alternation, Hume (2004)'s theory cannot account for why this occurs.[20]

Another more general problem is the fact that Hume (2004) cannot account for certain characteristics of metathesis, like the unidirectionality of some metatheses, how compensatory

[20] These results also crucially demonstrate sibilant metathesis in Modern Hebrew is not a Syllable Contact Law (SCL; Vennemann 1988) phenomenon. According to Vennemann (1988), there is a crosslinguistic preference for a falling sonority profile across a heterosyllabic consonant sequence (i.e. the more sonorant consonant should occupy the coda, as is the case with *hitpa'el* metathesis). However, if this were an SCL phenomenon, instead of a universal auditory phenomenon, as I argue below, the metathesis would be expected to affect all t + sonorant sequences, which it does not.

and coarticulatory metatheses occur, or why only sibilants, among fricatives, are predominantly involved in metathesis. Blevins & Garrett (2004) point out that the metathesis patterns PK > KP and T{P,K} > {P,K}T are unidirectional; they could not find any extant cases of the opposite patterns, accounting for this unidirectionality by noting that these stop-stop metatheses occur because of extreme gestural overlap. This extreme gestural overlap only occurs in coronal-noncoronal or labial-velar sequences because of the occurrence of regressive assimilation, which is only perceptually salient in coronal-noncoronal and labial-velar sequences (Blevins & Garrett 2004). When a coronal and noncoronal closure occur simultaneously, listeners often perceive the noncoronal closure first and, indeed, speakers move toward the noncoronal closure before the coronal closure begins (Blevins & Garrett 2004). Hume (2004) cannot explain these patterns except to say that noncoronal-coronal and velar-labial sequences must be more common in the world's languages, but this explanation does not clarify *why* this should happen. In fact, Hume (2004) claims that "in principle, any order of two segments is a potential output of metathesis, provided that the reordered sequence forms an attested structure in the language" (221). This is a bold claim which does not appear to be borne out by crosslinguistic research (including research by Hume herself), which demonstrates that only certain kinds of sounds are involved in metathesis, i.e. those for which there is some kind of phonetic motivation which causes ambiguity in the speech signal. Only Blevins & Garrett (2004) can account for the directionality of patterns of metathesis, by appealing to these specific phonetic motivations which engender metathesis.

Admittedly, the results from this experiment contradict Blevins & Garrett (2004)'s claim of the unidirectionality of TK > KT metathesis, as English speakers misperceived /kt/ as /tk/ the greater percentage of the time. Blevins & Garrett (2004) theorize that KT > TK metathesis does

not happen because regressive assimilation simply is not perceptually salient in noncoronal-coronal sequences; instead, they expect to find perceptual reanalysis of /kt/ to /kk/, /tt/, /k/, or /t/. It is conceivable that English-speaking listeners may have misperceived an ambiguous /kt/ sequence as something like /kk/, /tt/, /t/, or /k/, but, given the constraints imposed by the experimental task, which required listeners to choose one of two possible responses (a metathesized or unmetathesized stimulus verb), they could only choose /tk/ as the response most similar to what they actually misperceived. Such a scenario would explain why this particular misperception violates the presumed unidirectionality of TK > KT metathesis.

Blevins & Garrett (2004) also incorporate a wider range of phonetic explanations into their theory, which is able to account for metatheses which do not occur because of perceptual similarity, masking of phonetic cues, or "'stretched out' features," the only factors that Hume (2004) takes into account (217). Hume (2004)'s model is meant to address only CC metathesis, so that compensatory metathesis, involving "extreme anticipatory coarticulation" of a following vowel in a V_1CV_2 sequence, is unaccounted for, but she does not account for CC metathesis resulting from coarticulation either (Blevins & Garrett 2004, 125). Blevins & Garrett (2004) note a number of PK > KP metatheses, which involve two stops with different place features. It is unclear how this type of metathesis, which is apparently very widespread in Austronesian languages, and similar stop-stop metatheses like it, can be accounted for in Hume (2004)'s model.

Hume (2004) argues that there are two possible phonetic pathways leading to metathesis: 1) "diminished perceptual salience involve[ing] either similar sounds and/or those where the phonetic cues to the identification of at least one of the sounds is masked"; and 2) "'stretched out' features" which lead to acoustic overlap and indeterminacy, as the time domain of these

kinds of segments extends over adjacent sounds (217). Hume (2004) specifically points to homorganicity as one of the causes of "diminished perceptual salience" which can lead to indeterminacy and, eventually, metathesis (217). However, none of these phonetic scenarios applies to PK > KP metathesis. The sounds involved are not homorganic or similar other than the fact that they are both stops, and it is important to note that this kind of coarticulatory stop-stop metathesis only occurs with coronal-noncoronal and labial-velar sequences and never with any other combinations of stops. Hume (2004) cannot explain this restriction because she does not include this phonetic mechanism in her model. Since both sounds are stops, there can be no masking involved, and there are no stretched out features, like rhoticity, laterality, or palatalization, for example, which could cause this type of metathesis either. As such, this lack of empirical coverage is a critical flaw of Hume (2004)'s model.

Hume (2004) also cannot account for why sibilants are predominantly involved in metathesis, to the exclusion of other non-sibilant fricatives. Apart from /h/, there do not seem to be many cases of non-sibilant fricative metathesis; Ultan (1978) only mentions cases involving non-sibilant fricatives and /l/, which Blevins & Garrett (2004) would attribute to the long time domain features of laterality, not to anything inherent in non-sibilant fricatives. Cases of sibilant metathesis, however, are widespread, with Ultan (1978) specifically commenting on their prevalence. Hume (2004) explains that auditory-stream segregation leads to "separate auditory continua or streams [being] created among similar auditory cues and remain[ing] perceptually separated without temporal cross-linking," connecting this with her idea about stretched out cues leading to indeterminacy of temporal resolution (220). She points to an example from English [ho], with the frication of the /h/ superimposed over the vowel because of the long time domain of the aspiration (Hume 2004). Auditory-stream segregation is the same mechanism that Blevins

& Garrett (2004) hypothesize is behind auditory metathesis involving sibilants, as they argue that the process of auditory-stream segregation can lead to auditory stream decoupling with sibilants, where the frication is dissociated from its original place in a sequence of sounds. It is unclear from Hume (2004)'s account, though, how sibilant noise can be interpreted as a stretched out cue; Blevins & Garrett (2004)'s examples of stretched out cues include rhoticity, laterality, rounding, palatalization, velarization, pharyngealization, laryngealization, aspiration, retroflexion, and nasalization, all phonetic features which can affect nearby sounds or can be realized as secondary articulations (123). How can sibilant metathesis possibly fit into this explanation?

For Blevins & Garrett (2004), stretched out phonetic features cause metathesis because listeners have difficulty teasing apart the overlapping sounds, while auditory metathesis occurs because listeners have difficulty integrating different speech streams. This must occur because of a different mechanism. Blevins & Garrett (2004) believe that auditory stream decoupling is that mechanism, arguing that "a number of studies suggest that, in consonant clusters containing sibilants, the sibilant noise somehow distracts the listener, leading to high confusion rates with respect to the linear order of segments (Bregman 1990). Specifically, there is a tendency to decouple sibilant noise from the rest of the speech stream, and this decoupling can result in dramatic misperceptions" (128).

This idea comes from work by Bregman (1990) on auditory scene analysis, the process by which humans are able to make sense of acoustic events in their environments, determining what in the sensory input is "telling us about the same environmental object or event" (3). Bregman (1990) argues that auditory scene analysis is analogous to visual perception, in that humans have to "put the right combination of sensory evidence together…to recognize what is

going on" (3). This putting together of acoustic data is crucial, because, as Bregman (1990) points out, "Normally in perception, emergent properties are accurate portrayals of the properties of the objects in our environment. However, if scene analysis processes fail, the emergent perceived shapes will not correspond to any environmental shapes. They will be entirely chimerical" (5). This last point is especially important, because Blevins & Garrett (2004) argue that auditory scene analysis processes *do* fail in instances of auditory metathesis, with listeners perceiving an auditory stimulus which does not, in fact, exist. The major problem that auditory scene analysis must contend with is the messiness of the real world; the auditory system takes in physical acoustic events consisting of waves in air molecules which resound against the ear drum and resonate against different parts of the coiled cochlea (Bregman 1990). This produces what amounts to a spectrogram in the human brain, but this spectrogram is not like the spectrograms produced in silent sound booths, which display readily interpretable acoustic information (Bregman 1990). Instead, they consist of many different types of environmental sounds which are superimposed over each other, and the brain must sift through this information to make something meaningful (Bregman 1990). It does so by generating auditory streams, which are perceptual representations of physical events: "If correct perceptual representations of the world are to be formed, the evidence must be partitioned appropriately" (Bregman 1990, 6).

Bregman (1990) also argues that auditory perceptual organization is hierarchical, using the example of two types of instruments played together simultaneously. A human can understand that this acoustic event consists of two parts which make up a harmonious whole (Bregman 1990). He goes on to say that "raw organization into basic streams is the job of the perceptual system," while "the job of putting these streams into a hierarchical structure belongs to some more conceptual faculty of mind" (Bregman 1990, 204-205). In auditory stream

decoupling, the sibilant noise is segregated from the rest of the acoustic signal, so that when it comes time for the "more conceptual faculty of mind" (perhaps we might call this the phonology) to operate, it integrates the sibilant noise back into mental representation in a place where it did not exist previously, when the speaker originally pronounced it (Bregman 1990, 204-205). Schematically, this might look something like the following:

(36) Auditory stream decoupling:

S

/hit + SVCVC/ → hit VCVC → [hiStVCVC]

This is a plausible explanation of why sibilants are so often involved in metathesis, while other fricatives are not, unless they metathesize with a sound that has a long phonetic feature. Only Blevins & Garrett (2004) can explain this process, while Hume (2004) cannot, and conflating perceptual and auditory metathesis, when they are typologically distinct, is not theoretically sound.

6.2 The Bigger Picture: Metathesis and Sound Change

Metathesis can provide interesting evidence for theories of sound change which focus on sound change as phonetically motivated, as Blevins (2004) does. She convincingly argues that "the majority of commonly attested sound changes in the world's languages are mirrored by synchronic alternations of precisely the same type," sound changes and synchronic alternations being "nearly coextensive" (Blevins 2004, 4). Based on this fact, Blevins (2004) contends that "synchronic sound patterns are a direct reflection of their diachronic origins" and "that regular phonetically based sound change is the common source of recurrent sound patterns" (5). Thus, when a diachronic explanation exists, it should take precedence, to the exclusion of any

synchronic account, unless there is convincing evidence that a synchronic account is necessary (Blevins 2004). Essentially, Blevins (2004)'s theory of sound change, which she calls Evolutionary Phonology because of its emphasis on how synchronic sound patterns evolve from diachronic phonetic changes, is very similar to the Neogrammarian theory of sound change, except that she tries to explain even phenomena that the Neogrammarians excluded from their definition of sound change, because they believed that it was due to non-phonetic factors. This is especially evident from Blevins & Garrett (1998, 2004), which both successfully treat regular metathesis as the result of different phonetic factors, rather than a "trash can" category that comes about because of various non-phonetic reasons. Blevins (2004) does note that some sound changes arise through non-phonetic means, but she endeavors to try first to find a phonetic explanation before rejecting phonetics and appealing to non-phonetic explanations, like analogy or language contact.

Blevins (2004)'s account integrates typological work and experimental phonetic studies into her larger account of sound change, which allows her to argue that the typology of sound patterns in the world's languages is a direct result of the availability of phonetic pathways for change; the rarer patterns simply have no "common pathway of change which will result in their evolution" (Blevins 2004, 9). Importantly, like Ohala (1993), Blevins (2004)'s model of sound change is a listener-oriented model, meaning that the site of change exists within the listener. Listeners only have access to the acoustic signal, which may be ambiguous or easily misperceivable, leading to misconstrual of the speaker's intended speech form (Blevins 2004). If this happens often enough, a sound change will have taken place, with younger generations learning this modified pronunciation. Such a view of speech perception is in line with auditorist work (Halle & Stevens 1962; Stevens & Blumstein 1978, 1981) and work on modulation models

of speech perception (e.g. Story & Bunton 2010), which argue that humans process speech solely by reference to acoustics (with some visual information available, cf. the McGurk effect); this is in contradistinction to speech perception theories (e.g. the motor theory of speech perception, Liberman, Cooper, Harris, & Macneilage 1962) that hypothesize that listeners attend to a speaker's intended speech gestures, which are recoverable from the speech signal by some internal process. An auditorist model is simpler than a gestural model, as listeners directly experience the acoustic event and are able to perceive it, so it is preferable.

Blevins (2004)'s model of sound change is typological, with three possible pathways for sound change to take place. She refers to these as CHANGE, CHANCE, and CHOICE and calls her model the "CCC-model" (Blevins 2004, 32). CHANGE occurs when listeners mishear a sound as something perceptually similar, as when a speaker says [anpa], but a listener perceives [ampa] (Blevins 2004, 32). CHANCE occurs when a listener perceives an inherently ambiguous speech signal correctly, but interprets the signal in a different way from what the speaker intended (Blevins 2004). The example that Blevins (2004) cites involves a speaker saying [ʔa̰ʔ] for /aʔ/, while a listeners hears [ʔa̰ʔ] but presumes that the speaker intends to say /ʔa/. CHOICE involves phonetic variation, with a listener perceiving different variants correctly but "(a) acquir[ing] a prototype or best exemplar of a phonetic category which differs from that of the speaker; and/or (b) associate[ing] a phonological form with the set of variants which differs from the phonological form in the speaker's grammar" (Blevins 2004, 33). So, while a speaker may say [kakáta], [kăkáta], and [kkáta] for /kakata/, a listener may hear [kkáta], [kăkáta], and [kakáta] and presume that the speaker intends /kkata/ (Blevins 2004, 33).

In reference to CHANCE, Blevins (2004) notes that some "regular metatheses" are "reanalyses" of a "signal [that] is inherently ambiguous, though independent structural features

of the language may give rise to higher probabilities for one phonological analysis than another" (35). Sibilant metathesis, then, is an instance of CHANCE resulting not from a "language-independent phonetic bias" but language-specific tendencies to reinterpret ambiguous sequences (Blevins 2004, 35). She specifically mentions in a footnote that "Difficulties in feature localization may also arise as a result of auditory decoupling which may occur in the perception of sibilant and clock noise (Blevins and Garrett, to appear [2004]) (Blevins 2004, 35). In other words, while auditory stream decoupling is a failure of the auditory system which is not unique to Hebrew, the conditions in the language were right to induce metathesis in ambiguous stop-sibilant sequences. Experiment 2 at least partially replicates these original conditions, whatever they were, and induces misperception in English speakers which mimics Modern Hebrew metathesis.[21] The results of the experiment demonstrate that this phenomenon is not unique to Modern Hebrew, though language-specific conditions allowed it to occur, and that the phenomenon is not due to a reinterpretation of ambiguous sequences in accordance with the language's more common phonotactics, as Hume (2004) argues. The original cause of stop-sibilant metathesis in Modern Hebrew in particular may never be known, but this cannot be known for sure without continued research on sibilants, stop-sibilant clusters, and auditory decoupling. However, Blevins (2004) does mention that assymetries observed in metathesis and dissimilation patterns may be due to the phonologization of fast-speech phenomena, with articulatory timing playing a major role.

How might auditory stream decoupling have led to a situation where stop-sibilant metathesis is categorical within *binyan hitpa'el*? The answer lies in the process of

[21] It is important to note that Blevins (2004) argues that no misperception is actually involved, even though I continue to use the term in a colloquial sense. Blevins (2004) argues that listeners hear the ambiguous sequences correctly, but that they reinterpret their phonological instantiations.

phonologization. According to Hyman (2008), phonologization is the process whereby universal, intrinsic phonetic properties of speech become extrinsic and structured as part of a language's phonology; the phenomenon is one aspect of the larger phenomenon of grammaticalization. The inherently ambiguous nature of sibilants is such that historical speakers of Hebrew,[22] due to auditory-stream decoupling, misperceived the order of stop + sibilant sequences and posited a different underlying representation from the one the speaker intended. Once enough speakers began using the metathesized forms (probably due to a new generation misinterpreting the speech of an older generation), the metathesized variant became conventionalized and underwent phonologization to become a categorical part of the grammar. Even new *hitpa'el* verbs undergo metathesis, despite the fact that MH now has affricates, which might otherwise serve in these forms. As Hyman (2008) notes, the trajectory to phonologization follows the pattern of 1) "universal phonetics" (inherent ambiguity in stop-sibilant clusters) to 2) "language-specific phonetics" (speakers reanalyzing sibilant ambiguity as metathesized forms, at least some of the time) and, finally, to 3) "phonology" (categorical metathesis in *binyan hitpa'el*) (385).

7. Conclusion

Based on evidence provided by a speech perception experiment conducted with English speakers, I have demonstrated that metathesis of stop-sibilant clusters in Modern Hebrew cannot be the result of language-specific phonotactics. English-speaking listeners, when faced with degraded verbs in *binyan hitpa'el*, misperceived un-metathesized stop-sibilant clusters as sibilant-stop clusters, the same direction of misperception as occurs in Modern Hebrew metathesis. This was particularly relevant with the cluster [dz], which was misperceived at a

[22] Metathesis in certain verbal paradigms is a process which most Semitic languages share, so this particular phonologization may go as far back as Proto-Semitic.

higher rate as [zd] than [zd] was misperceived as [dz]. [zd], however, is the rarer of the two attested sequences in Modern Hebrew, a fact which contradicts Hume (2004)'s theory that indeterminate sound sequences are metathesized into the more common attested sequences in the language, because listeners use knowledge of their native languages to "split the difference" with indeterminate sounds and "guess" that the more common sequence is the one intended by the speaker. Importantly, this did not occur with the English speakers in the experiment.

Because Hume (2004)'s model cannot predict the outcome of these ambiguous sequences, suggesting that language-specific phonotactics are not at play, at least in the case of sibilants, I argue that Blevins & Garrett (2004)'s evolutionary phonology theory is the proper way to account for sibilant metathesis in Modern Hebrew. They argue, contrary to Hume (2004), that metathesis involving sibilants is a typologically distinct metathesis unlike the more predominant perceptual metathesis (which is what Hume (2004) believes all metatheses are) (Blevins & Garrett 2004). Instead, sibilant metathesis is an auditory metathesis, the outcome of a phenomenon called auditory stream decoupling, where the sibilant noise is separated from the rest of the auditory stream by the auditory system (Blevins & Garrett 2004). This is an occasional failure of auditory scene analysis, the process by which humans perceive the disparate sounds in their environments, and which is a part of the human auditory system in general (i.e. it is not a language-specific phenomenon) (Bregman 1990).

As Blevins (2004) convincingly argues, auditory metathesis is not the result of language-specific phonotactics, but rather a result of the universal process of auditory stream decoupling, which can induce metathesis in languages which have the specific acoustic characteristics that allow it to happen. What these acoustic characteristics are, I leave to future research on sibilants and stop-sibilant clusters, but it must rely on something that I have captured in the experiment,

because I was able to induce the same misperception in English speakers. Future research on other stop + sibilant sequences, like /k/ + sibilant and /p/ + sibilant, may shed more light on the phenomenon. Ravid (1995) notes that *hitpa'el* metathesis is acquired very early by Modern Hebrew-speaking children, with no reanalysis or "deviant phenomena" in "any section of the population"; she argues this is because the "morphophonemic rule" deriving *hitpa'el* metathesis is very transparent, only operating in this one instance (80). This may also suggest that the language-specific factors precipitating auditory stream decoupling are very strong in Modern Hebrew, and further research on L2 errors with *hitpa'el* metathesis, if they exist, would confirm this hypothesis.

Ultimately, this research demonstrates the fruitfulness of an approach to sound change which investigates changes in their possible phonetic motivating factors. In this way, phonologists can begin to explain previously unclear phonological phenomenon. The Neogrammarians, for example, believed that metathesis was not a phonetic change at all, but further research by later linguists has demonstrated beyond a reasonable doubt that metathesis is a regular process and one that is amenable to phonetic accounts, once its typology is taken into account (Blevins 2004). However, with advances in research on regular metathesis, we should not forget about sporadic metathesis, which remains a "trash can" category (Blust 2012). Phonologists should continue to investigate these curious cases because they explained in some motivated way, or they may represent unusual pathways of change, which are important to our understanding of the complexity of natural language (Blust 2005).

References

Amir Coffin, E., & Bolozky, S. (2005). *A reference grammar of Modern Hebrew*. Oxford: Oxford University Press.

Baden, J. S. (2010). Hithpael and niphal in Biblical Hebrew: Semantic and morphological overlap. *Vetus Testamentum, 60*(1), 33-44.

Bat-El, O. (1989). *Phonology and word structure in Modern Hebrew*. (Unpublished doctoral dissertation). University of California, Los Angeles. Available from ProQuest Dissertations and Theses database. (UMI No. 8912841).

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory*, *12*, 571-593.

Bat-El, O. (2003). The fate of the consonantal root and the binyan in Optimality Theory. *Recherches Linguistiques de Vincennes, 32*, 31-60.

Berman, R. A. (1978). *Modern Hebrew structure*. Tel Aviv: University Publishing Projects, Ltd.

Bickel, B. (2007). Typology in the 21st century: Major current developments. *Linguistic Typology, 11*(1), 239-251.

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK: Cambridge University Press.

Blevins, J., & Garrett, A. (1998). The origins of consonant-vowel metathesis. *Language, 74*(3), 508-556.

Blevins, J., & Garrett, A. (2004). The evolution of metathesis. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology*. Cambridge, UK: Cambridge University Press. 117-156.

Blust, R. (1996). The Neogrammarian hypothesis and pandemic irregularity. In M. Durie & M. Ross (Eds.), *The comparative method review: Regularity and irregularity in language change*. New York: Oxford University Press. 135-156.

Blust, R. (2005). Must sound change be linguistically motivated? *Diachronica, 22*(2), 219-169.

Blust, R. (2012). Hawu vowel metathesis. *Oceanic Linguistics, 51*(1), 207-233.

Boermsa, P., & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.19, retrieved 13 June 2016 from http://www.praat.org/

Bolozky, S. (1978). Some aspects of Modern Hebrew phonology. In R. A. Berman, *Modern Hebrew structure*. Tel Aviv: University Publishing Projects, Ltd. 11-67.

Bolozky, S. (1980). On the monophonematic interpretation of Modern Hebrew affricates. *Linguistic Inquiry, 11*(4), 793-799.

Bolozky, S. (1994). *The schizoid nature of Modern Hebrew: A Slavic language in search of a Semitic past*. Wiesbaden: Otto Harassowitz, 1990. [Review of Wexler 1990]. In R. A. Stone & W. P. Zenner (Eds.), *Critical essays on Israeli social issues and scholarship: Books on Israel, Vol. III*. 63-85.

Bolozky, S. (2008). *501 Hebrew verbs*. Hauppauge, NY: Barron's Educational Series.

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Butler, L. (2013). *Prosodically driven metathesis in Mutsun*. (Unpublished doctoral dissertation). University of Arizona.

Chayen, M. J. (1973). *The phonetics of Modern Hebrew*. The Hague, NL: Mouton.

Chomsky, N., & Halle, M. (1968). *The Sound Patterns of English.*

Coetzee, A. (1999). Metathesis in Tiberian Hebrew—A perspective from Optimality Theory. *Theoretical Linguistics, 25,* 99-131.

Dekel, N. (2014). *Colloquial Israeli Hebrew: A corpus-based survey*. Berlin: de Gruyter Mouton.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124.

Gafos, A. (2003). Greenberg's asymmetry in Arabic: A consequence of stems in paradigms. *Language, 79*(2), 317-357.

Greenberg, J. H. (1963). Afroasiatic. In J. Greenberg (Ed.), *The languages of Africa*. Bloomington, IN: Indiana University Press. 42-65.

Hale, M. (2003). Neogrammarian sound change. In B. D. Joseph & R. D. Janda (Eds.), *The handbook of historical linguistics*. Malden, MA: Blackwell Publishing.

Halle, M., & Stevens, K. N. (1962). Speech recognition: A model and a program for research.

Hayward, R. J. (2000). Afroasiatic. In B. Heine & D. Nurse (Eds.), *African languages: An introduction*. Cambridge, UK: Cambridge University Press. 74-98.

Hetzron, R. (1987). Hebrew. In Bernard Comrie (Ed.), *The world's major languages*. Oxford: Oxford University Press. 686-704.

Hock, H. H. (1985). Regular metathesis. *Linguistics, 23*, 529-546.

Hogg, R. M. (1977). Old English r-metathesis and generative phonology. *Journal of Linguistics, 13*(2), 165-175.

Horwood, G. (2004). Precedence faithfulness governs morpheme position. In L. Mikkelson & C. Potts (Eds.), *Proceedings of the 21st West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press.

Hume, E. (n.d.). Online metathesis database. Retrieved from http://metathesisinlanguage.osu.edu/default.cfm

Hume, E. (1998). Metathesis in phonological theory: The case of Leti. *Lingua, 104*, 147-186.

Hume, E. (2001). Metathesis: Formal and functional considerations. In E. Hume, N. Smith, & J. van de Weijer (Eds.), *Surface syllable structure and segment sequencing*. Leiden, NL: HIL. 1-25.

Hume, E. (2004). The indeterminacy/attestation model of metathesis. *Language, 80*(2), 203-237.

Hyman, L. (2008). Enlarging the scope of phonologization. UC Berkeley Phonology Lab Annual Report.

Jones, K. S. (2015). The phonetics of sibilant metathesis in Modern Hebrew. Unpublished manuscript.

Kiparsky, P. (1967). Sonorant clusters in Greek. *Language, 43*(3), 619-635.

Kutscher, E. Y. (1982). *A history of the Hebrew language*. Jerusalem: The Magnes Press.

Laufer, A. (1999). Hebrew. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association: A Guide to Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press. 96-99.

Lewis, M. P., Simons, G. F., & Fenning, C. D. (Eds.) (2016). *Ethnologue: Languages of the world, Nineteenth edition*. Dallas, TX: SIL International. Retrieved from http://www.ethnologue.com

Liberman, A. M., Cooper, F. S., Harris, K. S., & Macneilage, P. F. (1962). A motor theory of speech perception. Presented at the Speech Communication Seminar, Stockholm, 1-12.

Makashay, M. J. (2001). Lexical effects in the perception of obstruent ordering. In E. Hume, N. Smith, & J. van de Weijer (Eds.), *Surface syllable structure and segment sequencing*. Leiden, NL: HIL. 117-134.

Malone, J. (1993). *Tiberian Hebrew phonology*. Winona Lake, IN: Eisenbrauns.

McCarthy, J. J. (1979). *Formal problems in Semitic phonology and morphology*. (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

McCarthy, J. J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry, 12*, 373-418.

Even-Shoshan, A. (2003). *Milon Even-Shoshan (Even-Shoshan Dictionary)*. Jerusalem: Hamilon Hexadash, Inc.

Noonan, M. (2010). Genetic classification and language contact. In R. Hickey (Ed.), *The handbook of language contact* (48-65). Oxford: Blackwell Publishing.

Ohala, J. J. (1986). Introduction. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental phonology*. Orlando, FL: Academic Publishers. 1-12.

Ohala, J. J. (1987). Experimental phonology. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society 13*. 207-222.

Ohala, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives.* London: Longman. 237-278.

Pariente, I. (2010). Pharyngeal related non-lexical vowels in Sephardic Modern Hebrew. *Linguistics in Amsterdam, 3,* 1-19.

Prince, A. S. (1975). *The morphology and phonology of Tiberian Hebrew*. (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Prince, A. S., & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar.* Rutgers Cognitive Science Center Technical Reports.

Ratcliffe, R. (1997). Prosodic templates in a word-based morphological analysis of Arabic. In M. Eid & R. Ratcliffe (Eds.), *Perspectives on Arabic linguistics X*. Philadelphia: John Benjamins. 147-171.

Ravid, D. D. (1995). *Language change in child and adult Hebrew: A psycholinguistic perspective.* Oxford, UK: Oxford University Press.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/

Schwarzwald, O. (2005). Modern Hebrew consonant clusters. In D. D. Ravid & H. B. Shyldkrot (Eds.), *Perspectives on language and language development: Essays in honor of Ruth A. Berman*. Dordrecht, NL: Kluwer Academic Publishers. 45-60.

Smith, J. N. (2009). Source similarity in loanword adaptation: Correspondence Theory and the posited source-language representation. In S. Parker (Ed.), *Phonological argumentation: Essays on evidence and motivation* (pp. 155-177). London: Equinox.

Stevens, K., & Blumstein, S. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America, 64*(5), 1358-1368.

Stevens, K., & Blumstein, S. (1981). The search for invariant acoustic correlates of phonetic features. In P. D. Eimas & J. Miller (Eds.), *Perspectives on the study of speech*. Hillsdale, NJ: Erlbaum.

Story, B., & Bunton, K. (2010). Relation of vocal tract shape, formant transitions, and stop consonant identification. *Journal of Speech, Language, and Hearing Research, 53*, 1514-1528.

Styler, W. (2013). Using Praat for linguistic research. [Web document]. Version 1.3.5. Retrieved from https://phonetique.uqam.ca/upload/files/LIN2623/Styler_2013_2.pdf

Thompson, L. C., & Thompson, M. T. (1969). Metathesis as a grammatical device. *International Journal of American Linguistics*, *35*(3), 213-219.

Thibodeau, L. (n.d.). FM System Recordings. Retrieved from http://www.utdallas.edu/~thib/fm_wav.html

Ultan, R. (1978). A typological view of metathesis. In J. H. Greenberg (Ed.), *Universals of human language*. Stanford, CA: Stanford University Press. 367-402.

Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence. *Phonology, 16*, 401-442.

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language & Linguistic Theory, 23*(1), 169-218.

Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin.* Berlin: Mouton de Gruyter.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers, 36*, 481-487.

Warner, N. (1998). The Old Japanese vowel system: Implications of speech perception. In D. J. Silva (Ed.), *Japanese Korean Linguistics, Vol. 8*. Stanford, CA: CSLI Publications. 387-400.

Webb, C. A. Y. (1974). Metathesis. (Unpublished doctoral dissertation). UT Austin.

Wexler, P. (1990). *The schizoid nature of Modern Hebrew: A Slavic language in search of a Semitic past*. Wiesbaden: Otto Harassowitz.

Wilson, M. D. (1988). The MRC Psycholinguistic Database: Machine readable dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers*, *20*(1), 6-11.

Yanagawa, M. (2003). Metathesis in Modern Hebrew: An analysis in articulatory phonology. In M. J. Sole, D. Recasens, & J. Romero (Eds.), *15th International Congress of Phonetic Sciences*. 1671-1674.

Zuckermann, G. (2003). *Language contact and lexical enrichment in Israeli Hebrew*. New York, NY: Palmgrave Macmillan.

Zuckermann, G. (2009). Hybridity versus revivability: Multiple causation, forms and patterns. *Journal of Language Contact, 2,* 40-67.