creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

문학석사 학위논문

Learning Bias of Phonological Alternation in Children Learning English

영어를 습득하는 아동의 음운 교체에 대한 학습 편향

2017년 8월

서울대학교 대학원

영어영문학과 영어학 전공

조 진 영

Learning Bias of Phonological Alternation in Children Learning English

지도 교수 이 재 영

이 논문을 문학석사 학위논문으로 제출함

2017년 4월

서울대학교 대학원

영어영문학과 영어학 전공

조 진 영

조진영의 문학석사 학위논문을 인준함

2017년 6월

위 원 장 ____손 창 용____ (인)

부위원장 ____권 혁 승____ (인)

위 원 ____이 재 영____ (인)

Abstract

Learning Bias of Phonological Alternation in Children Learning English

Jo, Jinyoung

Department of English Language and Literature

The Graduate School

Seoul National University

The present study aims to test the claim in McCarthy (1998) and Hayes (2004) that output-to-output faithfulness constraints are ranked high *a priori* and that they are gradually demoted in the hierarchy as children are increasingly exposed to alternating forms. A prediction ensuing from this claim is that children will initially favor non-alternation, but they will eventually produce alternating forms as adults do.

In the acquisition literature (Bernhardt & Stemberger, 1998), it was reported that a child learning American English did not apply flapping where adults do, in conformity with paradigm leveling. For instance, the child produced *si*[t]*ing* for *sitting*, since flapping would trigger [t]~[ɾ] alternation between the base *si*[t] and the suffixed form *si*[ɾ]*ing*. Based on this anecdotal observation, this study investigates how the child learners' probability of producing alternating forms changes over time. Crucially, I explored the potential influence of input data, i.e., child-directed speech produced by mothers. Given that intervocalic /t/ eligible for flapping is occasionally realized as [t] (Boersma & Hayes, 2001; Hong, 2009), I tested the possibility that children's production of *si*[t]*ing* might simply reflect their tendency to mimic the relative frequencies of phonetic variants observed in their caregivers' speech.

Using spontaneous speech data of six mother-child dyads (Demuth, Culbertson, & Alter, 2006), I examined how word-medial /t/ and /d/ in the flapping environment were realized in suffixed words such as *eating* (base: *eat*) and *hiding* (base: *hide*). The results showed that children produced suffixed words with variants that are faithful to the corresponding segment in the base, especially in earlier stages of learning. First, for word-medial /t/, it was found that derived words were realized with [t] more frequently than non-derived words. Importantly, the probability of producing [t] in derived words was significantly higher in children's speech than in mothers' speech,

indicating that the children's production of [t] was not a direct replication of the relative frequencies of phonetic variants in the input. Similar results were found for word-medial /d/. The dominance of [d] over [ɾ] in terms of frequency persisted for a longer period in derived words than in non-derived words, suggesting that the bias for non-alternation in addition to articulatory difficulties in producing flaps encouraged the probability of producing [d] in derived words. Since derived and non-derived words did not significantly differ in their probability of being realized with [d] in mothers' speech, it is unlikely that children mimicked the frequencies of the variants attested in the input. Thus, I conclude that children's production of variants that are phonologically similar to the corresponding segment in the base is due to high-ranked OO-faithfulness constraints at the initial state of learning.

OT constraints are established in order to explain the patterns found in children's production, with ID-OO(son) and ID-OO(voi) accounting for the children's tendency to model their pronunciation of suffixed words on that of base. Based on these constraints, the children's grammar was captured with maximum entropy grammar (Goldwater & Johnson, 2003).

Keywords: learning bias, alternation, American English Flapping

Student Number: 2015-20056

Table of Contents

List of Tables

List of Figures

1. Introduction

Language learners often exhibit biases in learning phonological alternations. Evidence for learning biases among adult learners has been provided in numerous experimental works (e.g., White, 2014; Wilson, 2006). However, few studies have investigated the biases that young learners have when acquiring patterns of phonological alternation. This might be partially because we have little evidence for when children master patterns of alternation. As pointed out in Hayes (2004), a large amount of variation is observed among different morphological processes. Some studies documented quite an early mastery of morphology. For instance, it was reported in Aksu-Koç and Slobin (1985) that 15-month-olds learning Turkish were able to use the accusative suffix [-a]~[-e] correctly. However, other studies suggested that morphology is learned somewhat later than this. For example, N. Smith (1973) showed that a English-learning child at age 2;2 was not able to produce plurals correctly. In Berko's (1958) wug-test, even 4-year-old children failed to generalize the patterns of alternation in English plural suffixes to new items. Another case showing relatively late morphophonemic acquisition is the non-concatenative morphology of Modern Hebrew (Berman, 1985); young learners of Modern Hebrew did not master the patterns of non-concatenative morphology until they were 4-5 years.

Exploring the biases children possess in morphophonemic learning is nevertheless crucial in understanding some aspects of phonological acquisition. More specifically, it might help us provide an explanation for why children systematically make a particular pattern of errors. For instance, children learning Spanish incorrectly generalize diphthongization found in the 1sg, 2sg, 3sg and 3pl to the rest of the forms in a verb paradigm, which are supposed to be realized with non-diphthongized vowels (Clahsen, Aveledo, & Roca, 2002). Similarly, child learners of German used the vowel of infinitive, 1sg, 1pl, 2pl and 3pl forms in place of 2sg and 3sg (Clahsen, Prüfert, Eisenbeiβ, & Cholin, 2002). As will be discussed throughout this thesis, these overregularization patterns observed in the speech of young learners can be attributed to their innate bias towards non-alternation, or paradigm leveling. We can thus achieve deeper understanding of why children produce certain patterns of systematic errors by investigating biases in acquiring alternation.

The present study investigates child learners' bias towards non-alternation, focusing on the case of American English Flapping. According to Bernhardt and Stemberger (1998), a child beginning to learn English did not apply flapping rule where adults do, for the sake of paradigm leveling; for instance, she produced *si*[t]*ing* (rather than *si*[ɾ]*ing*) for *sitting*, since flapping would trigger [t]~[ɾ] alternation between the base form *si*[t] and the suffixed form *si*[ɾ]*ing*. Based on longitudinal corpus data of six child learners of

American English, this thesis aims to explore how the children's preference for non-alternation changes over time. The results showed that the frequency of variants that are faithful to the corresponding segment in the base (unsuffixed) form, such as *si*[t]*ing*, decreases as a function of age, while the frequency of [ɾ] is shown to increase over time[1]. The finding that child learners in fact are biased against alternation, especially in the initial stage of learning, provides empirical support for the claim in McCarthy (1998) and Hayes (2004) that output-to-output faithfulness (OO-faithfulness) constraints banning alternation was initially high-ranked, but later demoted in the learning process.

Importantly, I also investigated the potential influence of the input data, i.e., mothers' production of /t/ and /d/ in speech directed to children. It was found that mothers produced non-alternating forms in the initial stage of learning, indicating that children's higher probability of producing non-alternating forms in those periods might at least partially be affected by the way their mothers speak. However, it was revealed that the young children's initial preference for non-alternation could not be fully explained by the influence of their mothers' speech. The probability of producing non-alternating forms was significantly higher in children's speech than in

[1] In this thesis, I assume that OO-faithfulness constraints require faithfulness to a base form (Benua, 1997; Kenstowicz, 1997), which is assumed to be the unaffixed form for present purposes.

mothers' speech, evidencing an independent role of an innate bias towards non-alternation in child speech.

The rest of the thesis is organized as follows. The next chapter reviews previous research on children's biases in learning alternation. Chapter 3 describes the method and procedure used in this study. Then the results of the corpus study will be presented in Chapter 4, which is followed by a formal analysis of the results in Chapter 5. In Chapter 6, I present the results of learning simulations employing maximum entropy (maxent) grammar (Goldwater & Johnson, 2003) to model the children's learning trajectories. Chapter 7 discusses remaining issues and concludes this thesis.

2. Previous Studies

In Optimality Theory (Prince & Smolensky, 1993), it has been assumed that OO-faithfulness constraints are ranked high *a priori* by children acquiring a language (Hayes, 2004; McCarthy, 1998). The claim is that OO-faithfulness constraints are undominated in the initial stage of learning, i.e., prior to receiving any input data, and they are gradually demoted in the hierarchy as children are increasingly exposed to alternating forms. A prediction ensuing from this claim is that children will initially favor non-alternating forms, as high-ranked OO-faithfulness constraints require phonological similarity between morphologically related forms, but they will eventually be led to produce alternating forms as adults do. Moreover, assuming that the demotion of a constraint is a gradual process, the rate of demotion is determined by the frequency of their violations that young learners encounter in the input. Therefore, if one alternation occurs more frequently than another, the learner is expected to produce it earlier in the learning process.

Some empirical evidence for the initial bias towards non-alternation has been provided in the literature, though the reports have been only anecdotal and sporadic in form. For instance, it was reported in Kazazis (1969) that a 4-year-old learning Modern Greek innovated the illegal sequence *[xe] in producing [ˈe**x**ete] ‘you-pl. have’, the adult form of which is [ˈe**ç**ete], on the model of [ˈe**x**o] ‘I have’. Considering that the child had never found *[xe] in adult speech (as it violates phonotactics of the language), the markedness constraint that bans the sequence should be undominated. However, an OO-faithfulness constraint relevant to [x]/[ç] distinction was ranked even higher than the markedness constraint, leading the child to establish a non-alternating paradigm. A similar case, as already mentioned above, involves American English Flapping; according to Bernhardt & Stemberger (1998), a child learning American English produced [t] or [d], in order to satisfy OO-faithfulness constraints, where adults normally produce [ɾ]. More specifically, the child produced [sɪ**t**ɪŋ] ‘*sitting*’, instead of [sɪ**ɾ**ɪŋ], to base its pronunciation on [sɪ**t**] ‘*sit*’. In the same way, *needed* was realized as [niː**d**əd] rather than [niː**ɾ**əd] on the model of [niː**d**]. We can infer that [t] and [d] from unsuffixed forms, [sɪt] and [niːd], generalized to the corresponding suffixed forms [sɪtɪŋ] and [niːdəd], respectively. In terms of OT constraints, the avoidance of flaps in [sɪtɪŋ] and [niːdəd] satisfies OO-faithfulness constraints for [voice] and [sonorant][2]. If the suffixed form of [si**t**] is realized as [sɪ**ɾ**ɪŋ], the [t]~[ɾ] alternation between the two output forms violates OO-faithfulness constraints for [voice] and [sonorant]. Likewise, when the suffixed form of [niː**d**] is produced as [niː**ɾ**əd], the [d]~[ɾ] alternation between the base and its derived

[2] Throughout this thesis, I assume that English flap has [+sonorant] and [+voice] features.

counterpart violates OO-faithfulness constraints for [sonorant]. The authors of Bernhardt and Stemberger (1998) thus interpret that the child's tendency to avoid producing flaps in the suffixed forms is due to high-ranking OO-faithfulness constraints.

While these studies do provide evidence for the young learners' bias towards non-alternation, they are not sufficient to support the claim in Hayes (2004) and McCarthy (1998) that predicts children's initial preference for non-alternation but an increased probability of producing alternating forms in later stages of learning. Most importantly, since very few studies investigated how the child learners' probability of producing (non-)alternating forms changes over time, we lack empirical evidence showing that they indeed demote OO-faithfulness constraints as they increasingly encounter alternating forms along the developmental stages. Furthermore, we should also check whether the observations made in previous studies, most of which are based on only one child, are generalizable to a meaningful number of children. For example, if the occurrence of [t] or [d] in the speech of the child reported in Bernhardt and Stemberger (1998) is attributed to how young children's grammar is constructed, but not to some other idiosyncratic characteristics of that particular child's speech, the same tendency should be attested in the speech of every English-learning child. In these respects, we need more extensive and systematic research on children's bias towards non-alternation.

Recently, Do (2013) conducted a series of experiments to investigate

Korean children's learning trajectories of alternation shown in Korean noun and verb paradigms and modeled the results with maxent grammar. To my knowledge, this is the only study that adequately tested the claim that OO-faithfulness constraints are ranked high *a priori* and that they are demoted during the learning process. Children aged between 4 and 8 participated, whose patterns of alternation were expected to differ from those of adults. This seems to be a very late acquisition of alternations, which is probably due to a high degree of complexity found in Korean noun and verb paradigms. The results demonstrated that the younger children inflected nouns and verbs in a way that is faithful to one specific part of the paradigm, therefore satisfying OO-faithfulness constraints.

The main purpose of this study is to examine if children's prior bias towards non-alternation becomes weaker in the learning process due to the demotion of initially high-ranked OO-faithfulness constraints as claimed in previous studies (Hayes, 2004; McCarthy, 1998). As mentioned in Chapter 1, the present study also explores the influence of input data that children receive from their caregivers, an issue that has been neglected in previous studies. Although flapping is a predominant phonological process in American English, some variability is found in adult speech where intervocalic /t/ eligible for flapping is realized as [t] (Boersma & Hayes, 2001; Hong, 2009). In this sense, we must consider the possibility that children's production of [t] for *sitting*, if any, might reflect their tendency to mimic the relative frequencies of each phonetic variant observed in the input. Put differently, it might be the case that young learners occasionally produce *si*[t]*ing* simply because their caregivers do so, rather than because they obey OO-faithfulness constraints that require phonological similarity between *sitting* and *sit*. Due to the possibility that caregivers might produce *si*[t]*ing* rather than *si*[ɾ]*ing* in order to inform the young learners of the fact that *sit* and *sitting* are morphologically related, it becomes even more necessary to test the alternative hypothesis, i.e., children's production of *si*[t]*ing* is influenced by the input they receive from their caregivers. Thus, the present study addresses two important questions: Does child speakers' probability of producing alternating forms increase over time? If so, is their initial preference for non-alternation enforced by undominated OO-faithfulness constraints or by the influence of input? By exploring the influence of input on the children's production of /t/ and /d/, which has been rarely investigated in previous studies, this study aims to test the OO-faithfulness constraint hypothesis more rigorously and provide an extensive discussion on the observation made in Bernhardt and Stemberger (1998).

3. Methods

3.1. Data

The data examined in this study are the Providence Corpus (Demuth, Culbertson, & Alter, 2006), which consist of audio and video recordings of spontaneous speech interactions between six mother-child dyads. All six children (three boys, three girls) were monolingual speakers of American English.[3] Recording started when each child began to produce their first words (between the ages of 0;11 and 1;4) and ended when they were around 3 years old. Recordings from four dyads were collected once in every two weeks; the other two had denser corpora, with weekly recordings for approximately a year.

Children's speech across all periods in the corpus was subject to analysis. In this way, the youngest age from which the data were collected was 0;11 and the oldest 4;0. It was necessary to examine if the children produced /t/ or /d/ as flaps from the very beginning of their learning process. As discussed in the previous chapter, the rate of demotion of a constraint is

---

[3] It is stated in the transcripts that four of the six mothers (Mother 2, 3, 4, 6) were speakers of Standard American English, and the other two (Mother 1, 5) spoke Massachusetts dialect, which is characterized by its non-rhoticity. Compared to mothers who spoke Standard American English, Mother 1 and 5 showed a higher probability of producing [ʔ] for /t/ in suffixed words such as *getting* and *putting*. In Chapter 4 in which attested variants of /t/ are reported, I primarily focus on Child 4 – Mother 4 pair; therefore, such effect of dialectal difference would not be crucial in testing the hypothesis of interest in this study.

determined by the frequency of its violations in the input. Since flapping in American English is a widely observed process (Eddington, 2007; Patterson & Connine, 2001) that is found both within a word (e.g., *water*) and across word boundaries (e.g., *a lot of*), children might already have demoted the OO-faithfulness constraints by the time they produced their first words. In this case, children should be able to realize /t/ and /d/ as flaps. Meanwhile, it was also necessary to examine whether they actively produced flaps in much later learning stages. The reasons were twofold. First, by examining the probability of producing flaps, we can check whether the OO-faithfulness constraints are demoted or not by that time. If children are more likely to produce flaps for /t/ and /d/ in later stages of learning, we can infer that OO-faithfulness constraints are demoted during the learning process. Second, while OO-faithfulness constraints might no longer be active in children's grammar at that time, their production of flaps might still be restricted by articulatory constraints. Findings from previous studies (e.g., Klein & Altman, 2002; Rimac & Smith, 1984) suggested that children may still find it difficult to produce flaps even at around age 4, the oldest age from which recordings were obtained in the corpus. Klein & Altman (2002) showed that children learn to produce flaps gradually and that children as old as 48-60 months still could not produce flaps in an adult-like manner. Rimac & Smith (1984) pointed out that the fast movement of the tongue involved in producing flaps might pose an articulatory challenge for young children since they speak more slowly than adults do and that even 8-year-olds are still learning to produce flaps. Therefore, children’s speech from the entire period in the corpus, i.e., from approximately 1 to 4 years of age, was analyzed.

3.2. Target Words

In order to select the target words to be examined, I first generated a list in which all the words produced by the six children were presented in descending order of frequency. Then I manually extracted from the list all the disyllabic words that contain /t/ or /d/ in flapping environment, i.e., word-medial /t/ or /d/ followed by an unstressed vowel with an optional /r/ between the preceding vowel and /t, d/ (Kahn, 1976; Zue & Laferriere, 1979). The words thus chosen were categorized into four types (see Table 1), i.e., (a) derived words with medial /t/ (e.g., *eating*, base: *eat*), (b) non-derived words with medial /t/ (e.g., *better*), (c) derived words with medial /d/ (e.g., *hiding*, base: *hide*), (d) non-derived words with medial /d/ (e.g., *ready*). Words of type (a) and (c) were examined to see if younger children produced variants that are faithful to the base form, e.g., *ea*[t]*ing* and *hi*[d]*ing*. Words of type (b) and (d) were examined for comparison, for which there is no base form where OO-faithfulness constraints can operate. More specifically, if the children's production of *hi*[d]*ing* is motivated purely by its phonological resemblance to *hi*[d]*e*, *ready* should be produced with flaps since there is no reason for them to be produced with [d]. However, if *hi*[d]*ing* arises simply due to articulatory difficulties in producing flaps, *hiding* and *ready* should be realized with a similar proportion of [d]. Yet another possibility is that both factors, i.e., maintaining phonological similarity to the base and articulatory constraints, affect the realization of /d/. In this case, [d] will occur in both *hiding* and *ready*, but its occurrence should be somehow more robust in *hiding*.

Table 1 The target words examined for realization of /t/ or /d/

VtV					
Base: O			**Base: X**		
(a)	CHI	MOT	**(b)**	CHI	MOT
eating	172	604	*better*	111	595
getting	110	501	*butter*	69	267
putting	106	324			
sitting	104	312			
total	492	1741	total	180	862
VdV					
Base: O			**Base: X**		
(c)	CHI	MOT	**(d)**	CHI	MOT
birdy	79	101	*ready*	183	1197
hiding	69	243			
reading	36	173			
total	184	517	total	183	1197

Note: The numbers in CHI (child) and MOT (mother) columns indicate frequency of each word.

For type (a) and (c), words with a reasonably high frequency were chosen for analysis, i.e., higher than 20. For type (b) and (d), an attempt was made to select items whose frequencies roughly matched those of words in (a) and (c), respectively. Previous studies suggested that frequency might affect the realization of word-medial /t/ and /d/ in such a way that flapping was predominant in high frequency words, whereas low frequency words showed less frequent occurrence of flaps (Hong, 2009; Kreidler, 1989; Patterson & Connine, 2001). However, due to a paucity of disyllabic words with medial /d/, the frequencies of words in (c) and (d) do not match closely.

As there were only a small number of type (c) words in the corpus, *birdy* was included as a target word despite its phonological and morphological dissimilarity from *hiding* and *reading*. Phonologically, /d/ in *birdy* is preceded by /r/ whereas /d/ in *hiding* and *reading* occurs in intervocalic position. Morphologically, *birdy* has a different suffix (i.e., diminutive suffix -*y*) than *reading* and *hiding*. It seems that such phonological and morphological difference did not lead to a significant difference in flapping rate; the probability of undergoing flapping for *birdy* (27.8%) was higher than *hiding* (13.6%) and lower than *reading* (42.6%). Given that the mean age of children at which *birdy* tokens were found in the corpus (804 days) was earlier than that of *reading* (820 days) and later than that of *hiding* (700 days), it seems that the rate of flapping is conditioned mainly by the age at which children produced the word; the effect of phonological or morphological context in which /d/ occurs in *birdy*, if any, does not seem to be as strong as the effect of the age of production. Therefore, the results obtained from each word will be aggregated and reported altogether.

3.3. Coding

Realization of word-medial /t/ or /d/ was coded for [t], [d], [ɾ] or some other variants. I relied on acoustic measurement to categorize each token as one of the three major phonetic variants of /t/ and /d/, i.e., [t], [d] or [ɾ]. To be measured, a token was required to show acoustic evidence of being realized as [t], [d] or [ɾ], indicated by a sudden decrease in amplitude during the closure interval, a release burst after the closure and an abrupt increase in amplitude. Tokens were discarded when the speech was unrecognizable due to a significant overlap with background noise or another speaker's speech, when the voice was not loud enough, or when the recording quality was poor due to various reasons.

For each of the 'candidate' tokens that were acoustically clean and expected to be coded as one of the three variants, I measured closure duration and the percentage of the closure interval evidencing voicing (often referred to as voice bar). Closure duration was used to distinguish stops from flaps, and the percentage of closure that is voiced was used to categorize stop tokens into [t] or [d].

The tokens were first coded either as flap or stop based on closure duration, as it is widely accepted that short closure duration is the most prominent feature for the flap (Rimac & Smith, 1984; Steriade, 2000; Zue & Laferriere, 1979). Considering that mean closure duration for flaps in adult speech is reported to be 25-30ms (e.g., Byrd, 1993; Zue & Laferriere, 1979), the tokens extracted from mothers' speech with a closure duration of 30ms or shorter were coded as flap. Tokens from mothers' speech that exhibited a closure duration of 50ms or longer were coded as stop, as the mean closure duration for English voiced stops has been reported to be 50-60ms (and voiceless stops somewhat longer; Byrd, 1993; Flege & Brown, 1982). It was rather ambiguous to judge whether the tokens with an intermediate closure duration, i.e., between 30 and 50ms, should be coded as flap or stop. These tokens were not coded and excluded from analysis.

The cutoff values for mothers' speech, i.e., 30ms for flap and 50ms for stop, were also employed when coding the tokens from children's speech. One might argue that different cutoff values must be used for children's and mothers' speech, since children's speaking rate is significantly slower than that of adults and flaps produced by children might display longer closure duration. However, consonantal duration is much less sensitive to speaking rate compared with vowel duration (Magen & Blumstein, 1993; Yang, 2011) and findings are mixed concerning the effect of speaking rate on closure duration (Yao, 2007). Moreover, in Klein, Altman, and Tate's (1998) study in which phonetically trained investigators were asked to judge whether /t/ or /d/ produced by children in flapping environment was realized as a flap or as a stop, tokens with a closure duration of 3-17ms were judged as flap 75% of the time and 23-37ms 61% of the time. When the duration was 43-57ms, only 50% (which seems to be the chance level) of the tokens were judged as flap. Although previous studies rarely examined closure duration of flaps produced by children (probably due to the fact that children's production of /t/ and /d/ in the flapping context is often phonetically [d] rather than [ɾ]), we can infer that flaps found in child speech have a similar range of closure duration as that of adults (i.e., 10-40ms; Zue and Laferriere, 1979), as pointed out in Song, Shattuck-Hufnagel, and Demuth (2015). For these reasons, the same cutoffs used in coding the tokens found in mothers' speech seemed also appropriate for use in coding the tokens from children's speech.

Next, the tokens that were coded as stop were further investigated to be categorized as either [t] or [d]. Since voiced vs. voiceless stops are clearly distinguished by voicing during the closure (Flege & Brown, 1982; Wardrip-Fruin, 1982), [t]-[d] distinction was made based on the percentage of closure interval showing voicing.[4] For stops in V́CV context in adult speech, the

---

[4] It may have been possible to distinguish between [t] and [d] based on release duration (analogous to VOT in word-initial stops) in adult speech, as adult speakers were found to make a clear voiceless-voiced distinction using this parameter. However, it is known that cross-linguistically, children acquire unaspirated stops earlier than any other stop categories and might not be able to use this acoustic cue in producing aspirated stops and voiced stops (Jakobson, 1968; Kewley-Port & Preston, 1974). Macken and Barton (1980) found that English-learning children do not differentiate VOTs in voiced and voiceless word-initial stops.

mean percentage of voicing during the closure interval is usually 80-90% for [b] and 10-30% for [p] (Flege & Brown, 1982; Hanson & Shattuck-Hufnagel, 2011). As /t/ and /d/ in V́CV environment usually undergo flapping and thus show full voicing (i.e., voicing throughout the entire closure interval), we must consult other types of voiceless and voiced stops in establishing the cutoff values. Thus, following previous studies mentioned above, when voicing was observed in less than 30% of the closure interval, the given token was coded as [t]; when voicing persisted for more than 80%, the token was coded as [d]. The tokens with in-between values, i.e., 30-80%, were not coded and excluded from analysis.

Cutoff values for children's speech differed slightly from those of adults. While the cutoff for [t] was consistent across speakers, i.e., 30% for both mothers and children, tokens with more than 70% voicing were coded as [d] for children's speech. The cutoff was set at a lower value for children's speech, as it was suggested in previous studies that the mean percentage of voicing during closure for voiced stops in children's speech is lower than that of adults (Hanson & Shattuck-Hufnagel, 2011; Koenig & Lucero, 2008; B. L. Smith, 1979).[5]

In sum, for both children and mothers' speech, tokens with a closure

---

[5] Since there were only four tokens that had medial voicing between 70 and 80% in children's speech, the effect of lowering the cutoff value for children's production should not be very substantial.

duration of 30ms or shorter were coded as flap, while those with a closure duration of 50ms or longer were coded as stop. For children's speech, a token was coded as [t] when it showed voicing in less than 30% of the closure interval, and [d] when more than 70%. In mothers' case, the cutoffs were 30% and 80% for [t] and [d], respectively.

As noted earlier, the tokens that were not acoustically clean enough to be examined were excluded from analysis. Thus, a substantial number of tokens were discarded due to less-than-ideal quality of the recordings; of the 5,356 tokens that were originally found in the corpus (see Table 1), only 3,715 were acoustically coded.[6] To verify that investigating only a part of the data did not affect the results in testing the hypothesis, I also used perceptual coding (as a second best method) and confirmed that the results did not crucially differ depending on the number of tokens investigated. The result of acoustic coding was consistent with that of perceptual coding in 87.6% of the total number of tokens for children's speech and 98.2% for mothers'. The results presented in the next chapter are based solely on acoustic coding.

---

[6] Note that because the quality of the recordings in the Providence corpus is not ideal, there are cases where researchers examine only the tokens that are acoustically clean in segmental analysis, e.g., Song et al. (2015) used only the first ten acoustically clean tokens of each target word. Still, this corpus contains a larger amount of data with a relatively good recording quality compared to other child speech corpora.

4. Results

In this chapter, I first present a descriptive acoustic analysis of word-medial /t/ and /d/ with respect to closure duration and voicing during the closure. Then the observed frequencies of the variants of /t/ and /d/ in children's speech are reported, followed by those of mothers' speech. The results for /t/ and /d/ will be presented separately, and within each consonantal condition, suffixed words (e.g., *eating*) and words with no base form (e.g., *better*) are compared in terms of how /t/ or /d/ were phonetically realized in each morphological context.

4.1. An Acoustic Analysis of Word-medial /t/ and /d/

Figure 1 shows the distribution of closure durations of all tokens across children's age observed in (a) children's and (b) mothers' speech. In line with findings from previous studies (e.g., Zue & Laferriere, 1979), closure durations obtained from mothers' production of /t/ and /d/ in flapping environment mostly ranged from 10 to 40ms. In comparison, children's production of /t/ and /d/ exhibited much longer closure durations, confirming the findings from the acquisition literature (Klein et al., 1998; Rimac & Smith, 1984). The two horizontal lines in each figure represent the cutoff values used in acoustic coding; the tokens above the upper line (y=50) were coded as stop, while those below the lower line (y=30) were coded as flap.

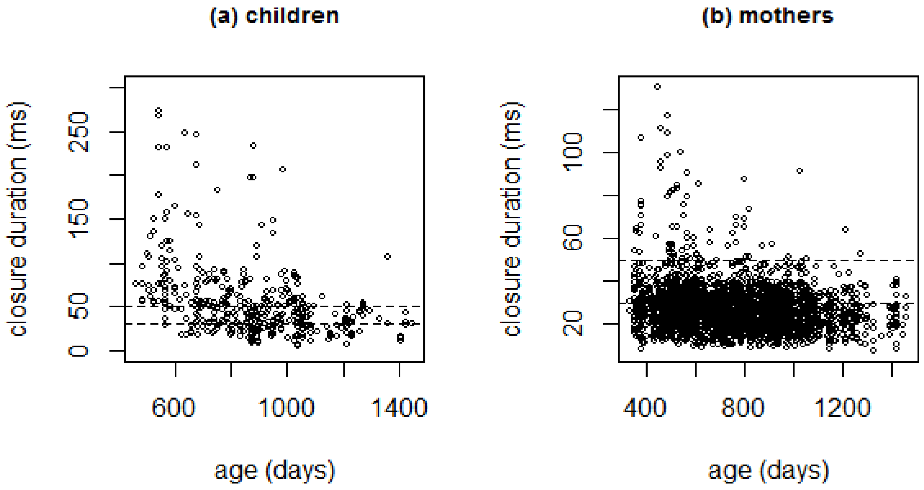


Figure 1 The distribution of closure duration of all tokens across children's age observed in (a) children's and (b) mothers' speech. (*Note*: Two outlier tokens are not presented in (a), which had an unusually long closure duration of around 600 ms.)

In Figure (1a), it is observed that the number of tokens with a closure duration of 50ms or longer decreased as children grew older, indicating that children were more likely to produce [t] and [d] variants when they were younger. This is in line with the prediction that children will favor non-alternation in the beginning stage of learning. Interestingly, a similar tendency was found in mothers' speech as well. It is shown in Figure (1b) that mothers' probability of producing [t] and [d] also decreased as a function of children's age. This indicates that mothers were more likely to produce non-alternating forms in the early stage of children's learning process. A more detailed description on the frequency of non-alternating forms in children's and mothers' speech will be provided in 4.2 and 4.3, respectively.

Figure 2 presents the percentage of closure interval evidencing voicing for tokens that were coded as stop in (a) children's and (b) mothers' speech. It seems that the proportion of closure interval evidencing voicing can be a reasonably good measure by which the tokens are categorized into [t] and [d]. Although there are quite a few tokens that fall between the two cutoff values (represented by the two horizontal lines) and thus cannot be judged as either [t] or [d], it seems that phonetic [d] shows dramatically high rates of full voicing (100%) and phonetic [t] shows very low percentage of voicing, as reported in previous research (Flege & Brown, 1982; Keating, 1984; Wardrip-Fruin, 1982).

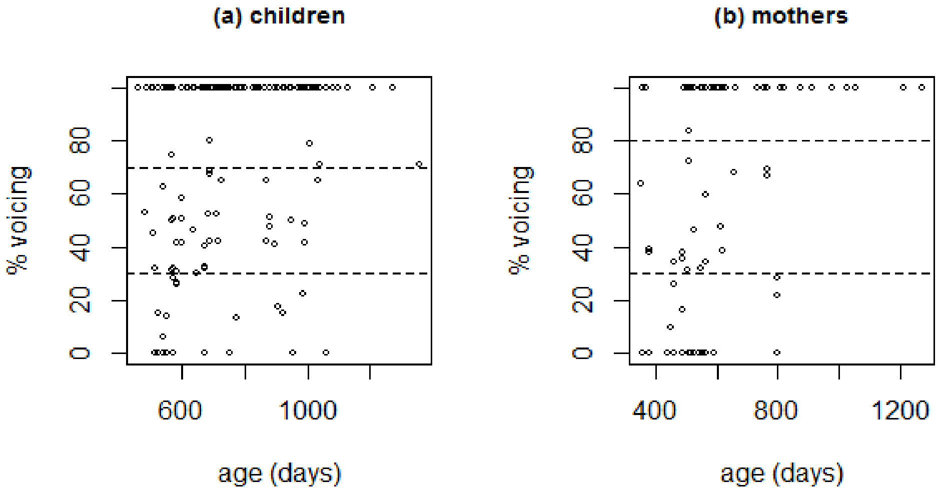


Figure 2 The distribution of percentages of voicing during the closure across children's age observed in stops produced by (a) children and (b) mothers

4.2. Children's Production of Word-medial /t/ and /d/

4.2.1. Realization of Word-medial /t/ in Child Speech

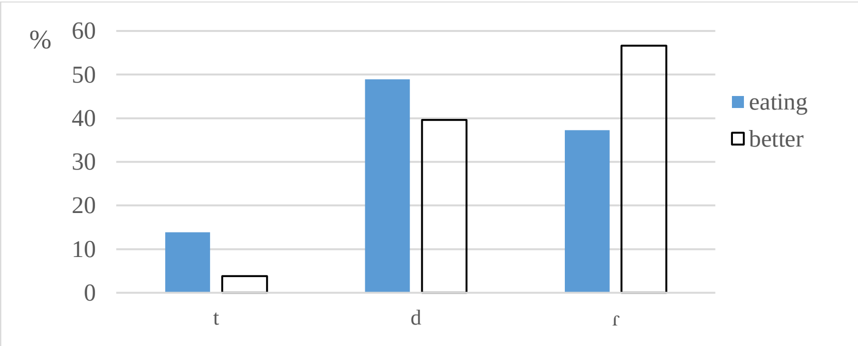
Throughout the remaining parts of Chapter 4, I present the frequencies of the phonetic variants of /t/ and /d/ in two steps, the latter of which is crucial in this study. I will first show overall results including all tokens investigated in this study, with the category "[t], [d], [ɾ] candidate" (see Table 2). The tokens belonging to this category showed acoustic evidence of being realized as [t], [d] or [ɾ], (e.g., an abrupt decrease in amplitude during the closure followed by a release burst) and were eligible for acoustic investigation for closure duration and voicing during the closure. For example, Table 2 shows that 357 tokens for *eating* words were acoustically clean and their phonetic realization could be determined, and that 50.7% of these tokens (n=181) were to be judged as one of the three variants [t], [d], or [ɾ]. Next, I focus on how these candidates are actually categorized, determined by the cutoff values established in Chapter 3. For instance, Table 3 shows that among the 181 candidate tokens, 17 were coded as [t], 59 as [d] and 46 as [ɾ]; the rest were excluded because they did not meet the cutoffs. This should be the crucial part of the results based on which the hypothesis concerning OO-faithfulness is tested.

Table 2 Attested variants of /t/ in *eating* and *better* words in child speech

	[t] [d] [ɾ] candidate	[Ø]	[n]	[ʔ]	others [7]	total
eating words	181 (50.7)	138 (38.7)	24 (6.7)	4 (1.1)	10 (2.8)	357 (100)
better words	81 (62.8)	34 (26.4)	0	0	14 (10.9)	129 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed.

Table 3 The number of tokens judged as [t], [d] and [ɾ] for /t/ in *eating* and *better* words in child speech



	[t]	[d]	[ɾ]	total
eating words	17 (13.9)	59 (48.9)	46 (37.3)	122 (100)
better words	2 (3.8)	21 (39.6)	30 (56.6)	53 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant.

[7] In "others" category for both *eating* and *better* words, [l] accounted for the majority of the variants, which was produced mostly by Child 1.

Table 3 presents the number of tokens coded as [t], [d] and [ɾ] for /t/ in derived words (*eating* words) and non-derived words (*better* words) produced by children. The first thing to examine was whether children actually produced [t] in *eating* words that is faithful to the base form (e.g., *ea*[t]). It was necessary to compare the probability of producing [t] in *eating* and *better* words in order to examine if the occurrence of [t] in *eating* words was indeed motivated by OO-faithfulness constraints. As the total number of tokens examined was different for *eating* and *better* words, comparisons were made in terms of the percentage of tokens coded as each variant, rather than the raw frequency of each variant.

While *eating* words were realized with [t] in 13.9% of the tokens, only 3.8% of the tokens for *better* words were realized with [t]. Assuming that the base forms such as *eat* are realized with [t], this result can be interpreted as reflecting the children's tendency to base the pronunciation of *eating* words on that of their base forms. In other words, since *better* words have no base form to resemble, the reason why [t] variants were attested almost exclusively in *eating* words is probably due to their phonological resemblance to their base forms, enforced by OO-faithfulness constraints.

However, an alternative explanation to the more frequent occurrence of [t] for *eating* words is that children produce [t] as an alternative pronunciation to [ɾ] at a very young age when their articulatory systems are immature, and that the age from which *eating* words were extracted was somewhat earlier than *better* words. Under this scenario, the reason why the children more frequently produced [t] in *eating* words would be simply that they happened to produce *eating* words in earlier periods than *better* words. Since the mean age (in days) at which *eating* and *better* words were produced was 841 and 928, respectively, and the difference was statistically significant (p<0.05), we must suspect that *better* words might not have been realized with [t] because children produced those words in later stages of learning. Therefore, for a more accurate comparison between the two morphological contexts, it was necessary to consider the age of production when we explore the changes in frequencies of each variant.

Figure 3 and Figure 4 display the distribution of [t], [d] and [ɾ] tokens across children's age in *eating* words and *better* words, respectively.

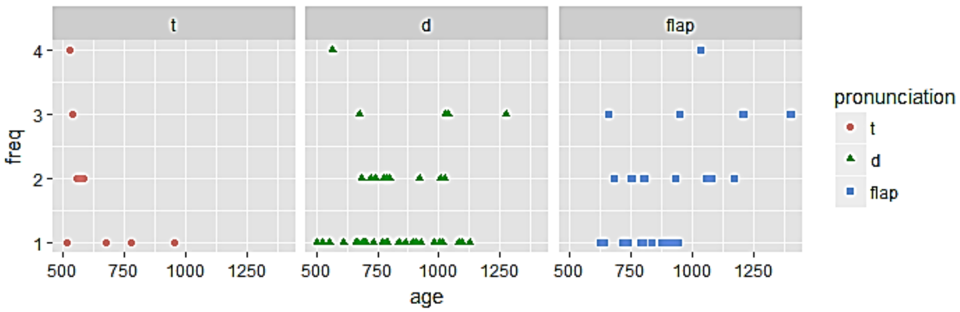


Figure 3 Distribution of [t], [d] and [ɾ] for *eating* words in children's speech across children's age (in days)

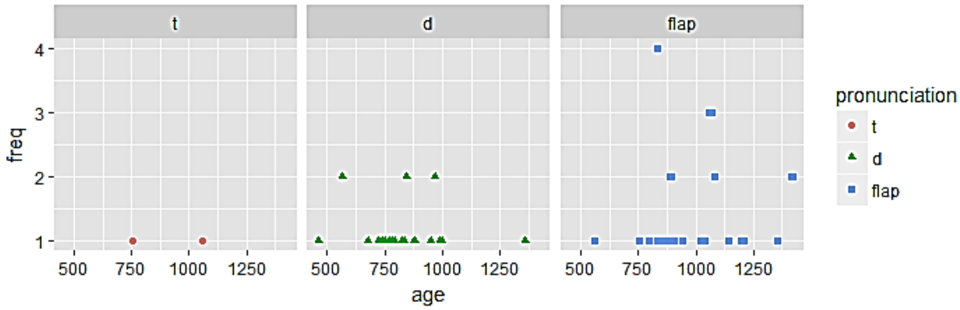


Figure 4 Distribution of [t], [d] and [ɾ] for *better* words in children's speech across children's age (in days)

We can observe in Figure 3 and Figure 4 that frequencies of each variant undergo some changes over time. First, [t] variants for *eating* words were mostly found in very early periods. Also, for both *eating* and *better* words, it seems that [d] variants were more frequently found in earlier periods, while [ɾ] was more likely to be produced in later periods. The picture becomes clearer when we divide the learning process into three distinct stages as in Table 4. I first divided the entire period into several sub-periods, each of which covers 50 days, in order to track the changes in frequencies of [t], [d] and [ɾ]. As major changes in children's production were observed around 800 and 1050 days for both /t/ and /d/ words, I grouped the sub-periods into three stages; before 800 (Early stage), between 800 and 1050 (Intermediate stage), and after 1050 (Later stage). Within each stage, the mean age from which the tokens were extracted did not significantly differ depending on the type of words, i.e., /t/ in derived words, /t/ in non-derived words, /d/ in derived words and /d/ in non-derived words.[8] Thus, a fair comparison could be made across different consonantal and morphological conditions.

Table 4 shows frequencies of each variant for *eating* and *better* words produced by the children in the three stages.

Table 4 Frequencies of [t], [d] and [ɾ] for *eating* and *better* words produced by children in three stages

		Early	Intermediate	Later
eating words	[t]	16 (27.1)	1 (2.2)	0
	[d]	30 (50.8)	23 (51.1)	6 (33.3)
	[ɾ]	13 (22.0)	21 (46.7)	12 (66.7)
	total	59 (100)	45 (100)	18 (100)
better words	[t]	1 (7.7)	0	1 (6.3)
	[d]	10 (76.9)	10 (41.7)	1 (6.3)
	[ɾ]	2 (15.4)	14 (58.3)	14 (87.5)
	total	13 (100)	24 (100)	16 (100)

Note: The numbers in parentheses indicate the percentage of tokens coded as each variant within each stage.

[8] One-way ANOVA was performed in each stage to compare the mean age at which the four types of words were produced. It was found that in the Early and the Later stage, the effect of word type on age of production was not significant (Early F(3, 95)=1.985, *p*=0.121; Later F(3, 49)=1.262, *p*=0.298). In the Intermediate stage, a significant main effect of word type was found (F(3, 118)=2.684, *p*=0.0499), but a post-hoc Tukey HSD test indicated that the difference did not reach a significant level.

It is shown that [t] variants were mostly produced before children reached 800 days of age. Crucially, while 16 tokens of [t] for *eating* words were found in the Early stage, only one [t] token was found for *better* words in the same period. The fact that *eating* words were more frequently produced with [t] compared to *better* words even when the age factor was controlled seems to suggest that children tend to base their pronunciation of *eating* words on that of the corresponding base forms such as *eat*. However, a Fisher's Exact Test revealed that the probability of producing [t] for *eating* words was not significantly different from that of *better* words (p=0.17), which might weaken the claim that OO-faithfulness constraints enforce phonological similarity between *eating* words and their base forms.

Nevertheless, there are several reasons to believe that the children produced *ea*[t]*ing* on the model of the base form *ea*[t]. First, when we look at individual speaker's production patterns as in Table 5, it turns out that most of the tokens were contributed by one child (Child 4). (Recall that Child 3 and Child 4 had denser corpora than the other four; since the tokens were excluded either because they were not clear enough to be acoustically investigated or because they did not meet the cutoff criteria, only a few token were found in the speech of Child 1, 2, 5 and 6.) Focusing on this one child, the contrast between *eating* words and *better* words becomes clearer.

Table 5 Frequencies of [t], [d] and [ɾ] for *eating* and *better* words produced by each child

		CHI 1	CHI 2	CHI 3	CHI 4	CHI 5	CHI 6
eating words	[t]	0	0	1	16	0	0
	[d]	5	5	11	36	2	0
	[ɾ]	0	0	6	40	0	0
better words	[t]	0	0	2	0	0	0
	[d]	0	2	3	14	1	1
	[ɾ]	3	0	5	15	2	5

Table 6 Frequencies of [t], [d] and [ɾ] for *eating* and *better* words produced by Child 4 in three stages

		Early	**Intermediate**	**Later**
eating words	[t]	15 (30.0)	1 (3.4)	0
	[d]	24 (48.0)	11 (37.9)	1 (7.7)
	[ɾ]	11 (22.0)	17 (58.6)	12 (92.3)
	total	50 (100)	29 (100)	13 (100)
better words	[t]	0	0	0
	[d]	9 (90.0)	5 (45.5)	0
	[ɾ]	1 (10.0)	6 (54.5)	8 (100.0)
	total	10 (100)	11 (100)	8 (100)

Note in Table 5 that [t] variant was not observed for *better* words in the speech of Child 4, while there were 16 [t] tokens for *eating* words. As already mentioned, most of these tokens were attested in the initial stage of learning as shown in Table 6. In the Early stage, she produced 15 tokens of [t] for *eating* words but none for *better* words, suggesting that this child was biased towards producing non-alternating forms in the initial stage of learning. A Fisher's Exact Test showed that the difference in the probability of producing [t] for *eating* words and *better* words in the Early stage was marginally significant (p=0.0535).

There is yet another piece of evidence that suggests children's production of *ea*[t]*ing* is modeled on the base form *ea*[t]. [t] variants for *eating* words were not only more frequently attested, but also qualitatively different from [t] variants found in *better* words. More specifically, [t] tokens in *eating* words were phonetically more similar to the way /t/ in *eat* words is typically realized, characterized by a clear closure and a release. The mean release duration (word-medial VOT) of [t] in *eating* words was 54ms, while that of [t] in *better* words was 10ms. The difference was statistically significant ($p<0.05$). One can thus say that only /t/ in *eating* words had a clear release and that compared with /t/ in *better* words it was acoustically more similar to the realization of /t/ in *eat* words. Such difference suggests that the occurrence of [t] tokens in *eating* words was motivated by its phonological similarity to the corresponding segment in base forms such as *eat*.

So far, I have maintained that children's production of [t] in suffixed words was modeled on their base forms, motivated by high-ranked OO-faithfulness constraints. An implicit assumption was that /t/ in the base forms is realized as [t]. In other words, the hypothesis concerning OO-faithfulness constraints can only be supported when we can verify, for example, that /t/ in *eating* and *eat* are phonologically similar to each other. I investigated the phonetic realizations of word-final /t/ in *eat* words.[9] As /t/ in word-final position can have many allophonic variants (e.g., unreleased stop, glottal stop and flap) depending on the type of segment that follows, it was necessary to examine its phonetic variants in an isolated context in which it was not followed by any segment. For example, if a child produced *ea*[ɾ] *it*, it is ambiguous whether the occurrence of the flap resulted from flapping rule in the relevant context (i.e., followed by an unstressed vowel) or the child actually considered [ɾ] as a phonetic variant of /t/ in word-final position. As a simple way of ruling out any contextual effect, only the tokens that occurred in utterance-final position were investigated. In this way, we can more accurately examine whether the occurrence of [t] observed in *eating* words was in fact based on [t] in *eat* words.

Table 7 presents how /t/ in *eat* words is realized in the children's speech.

---

[9] The coding scheme followed that of Song et al. (2015).

Table 7 Realization of word-final /t/ in *eat* words produced by children

	[t] [10]	Ø	others [11]	total
eat words	195 (85.9)	24 (10.6)	8 (3.5)	227 (100)

Note: The numbers in parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed.

It was found that more than 85% of word-final /t/ in *eat* words was realized as the alveolar stop in children's speech. Thus, one can argue that /t/ in the suffixed form such as *eating* is phonologically similar to /t/ in the base form, at least when the coda /t/ is not followed by any segment.

A more precise investigation should concern children's production of /t/ in the base form in the early stage of learning, since it is the beginning stage of learning when they were found to produce *ea*[t]*ing*. Table 8 shows phonetic variants of /t/ in *eat* words produced by Child 4 in each stage. We focus on this one specific child's production of the base forms, as generalizations on production of the suffixed forms were mostly based on her speech. As shown in Table 8, in the Early stage (during which she produced *ea*[t]*ing*), the majority of the *eat* words were produced with [t] (97.8%). In fact, the percentage of tokens coded as [t] was very high in all stages.

[10] This category includes the following phonetic variants: [t] with a clear release (72.2%), [t̚] (12.8%), [tV] (0.4%) and [dV] (0.4%), where [V] is an epenthetic vowel.
[11] In "others" category, tokens were included for which the child speakers produced fricatives for /t/ in *eat* words.

Therefore, phonological resemblance between a suffixed form and its base has been confirmed in the speech of Child 4.

Table 8 Realization of word-final /t/ in *eat* words produced by Child 4 in three stages

	Early	**Intermediate**	**Later**
[t]	87 (97.8)	22 (95.7)	6 (100)
others	2 (2.2)	1 (4.3)	0
total	89 (100)	23 (100)	6 (100)

Note: The numbers in parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed within each stage.

I have focused so far on the distribution of [t] tokens in children's speech and suggested that the more frequent occurrence of [t] variants in *eating* words, especially in early periods of learning, can be empirical evidence to the claim in McCarthy (1998) and Hayes (2004) that OO-faithfulness constraints are high-ranked *a priori* but demoted in the learning process. I provide further evidence to this claim in 4.2.2. by investigating children's production of word-medial /d/. Before that, I briefly comment on the distribution of [d] and [ɾ] in *eating* and *better* words.

The results of the coding revealed that the children tended to produce [d] more frequently in earlier days, while they produced [ɾ] more frequently in later stages (see Table 4). This confirms the findings from previous studies that flaps might be challenging for young learners to produce due to articulatory constraints and that they most often produce [d] where adults produce flaps (Klein & Altman, 2002; Rimac & Smith, 1984). Only when they become more experienced learners and have a better command of articulatory systems are they able to produce flaps actively. Frequencies of each variant are presented in (1), a simplified version of Table 4.

(1) Frequencies of each variant for *eating* and *better* words in child speech

		Early	**Intermediate**	**Later**
eating words	[t]	16	1	0
	[d]	30	23	6
	[ɾ]	13	21	12
	total	59	45	18
better words	[t]	1	0	1
	[d]	10	10	1
	[ɾ]	2	14	14
	total	13	24	16

For both *eating* and *better* words, the frequency of [d] outnumbered that of [ɾ] in the Early stage (*eating* 30 vs. 13; *better* 10 vs. 2). When they reached the Later stage, the pattern was reversed; [d] tokens were less frequently observed than [ɾ] tokens (*eating* 6 vs.12; *better* 1 vs. 14).

In the Intermediate stage, however, *eating* and *better* words exhibited different patterns. While [d] variants were more frequently found than [ɾ] for *eating* words (23 vs. 21), the opposite was true for *better* words (10 vs. 14). Such difference in the probability of producing [d] in the two morphological contexts may be due to the effect of OO-faithfulness constraint for [sonorant]. Consider first that [d] as a realization of /t/ in *eating* words violates OO-faithfulness constraint for [voice], while [ɾ] violates OO-faithfulness for both [voice] and [sonorant] as shown in (2).

(2) Violation of OO-faithfulness constraints by [t], [d] and [ɾ] variants

/iːtɪŋ/ base: [iːt]	ID-OO(son)	ID-OO(voi)
iːtɪŋ		
iːdɪŋ		*
iːɾɪŋ	*	*

/betəɹ/ base: none	ID-OO(son)	ID-OO(voi)
be.təɹ		
be.dəɹ		
be.ɾəɹ		

In this sense, [ɾ] is less harmonic than [d] when evaluated by OO-faithfulness constraints. However, this is not the case in *better* words; as all the candidates (vacuously) satisfy OO-faithfulness constraints, [d] and [ɾ] are on equal grounds with regard to these constraints. In other words, because *better* words have no morphologically related forms, there is no way that the OO-faithfulness constraints can operate to prefer [d] to [ɾ]. Assuming that articulatory constraints that enforce producing [d] (rather than [ɾ]) affect *eating* and *better* words to the same extent, the reason why the ratio of [d] to [ɾ] was higher in *eating* words in the Intermediate stage might be that [ɾ] in *eating* words violates an additional OO-faithfulness constraint, i.e., ID-OO(son), than [d] (which is not the case in *better* words.) Though the difference was not statistically significant (Pearson's Chi-squared test with Yates' continuity correction; $\chi^2=0.3392$, df=1, $p=0.56$), the results of the learning simulation in Chapter 6 seem to indicate that OO-faithfulness constraint for [sonorant] is at work in this learning stage, reflected in the weight of ID-OO(son).

If ID-OO(son) enforces a higher ratio of [d] to [ɾ] for *eating* words compared with *better* words in the Intermediate stage, why is such pattern absent from the Early stage? At first glance, it seems that the relative frequencies of phonetic variants in the Early stage do not conform to a pattern that is expected from the effect of ID-OO(son) explained above, as the ratio of [d] to [ɾ] seems to be lower in *eating* words (30 vs. 13) than in *better* words (10 vs. 2). This might be due to the small number of tokens examined in this study, especially for *better* words in the Early stage. When we combine the frequencies of each variant obtained from the first two stages as in (3), where

I assumed ID-OO(son) to be active, we can see that the frequencies now reflect the effect of ID-OO(son) in an expected way:

(3) Frequencies of variants in the Early and Intermediate stage in child speech

		Early - Intermediate
eating words	[t], [d]	70
	[ɾ]	34
	total	104
better words	[t], [d]	21
	[ɾ]	16
	total	37

Note that the ratio of the combined frequency of [t] and [d] to the frequency of [ɾ] is higher in *eating* words (70 vs. 34) than in *better* words (21 vs. 16), consistent with the prediction that the probability of producing variants that violate ID-OO(son) would be discouraged in the suffixed forms.

To summarize the findings for realization of /t/ in derived and non-derived words, it was found that [t] variant, which is maximally faithful to the base forms (e.g., *ea*[t]), was more frequently attested in derived words. I argue that this is due to OO-faithfulness constraints that require phonological identity between morphologically related words. One might wonder if such result is due to the influence of input; it could be the case that mothers also produced a small amount of [t] variants for some reason, and children simply mimicked the pattern observed in the input. I consider this possibility in section 4.3.1.

4.2.2. Realization of Word-medial /d/ in Child Speech

This section presents the results of the coding for the children's production of /d/ in suffixed words (*hiding* words) and in words with no base form (*ready* words). Table 9 presents the overall results of the coding including all tokens examined in this study. In both *hiding* and *ready* words, around 60% of the tokens were judged as either [t], [d] or [ɾ].

Table 9 Attested variants of /d/ in *hiding* and *ready* words in child speech

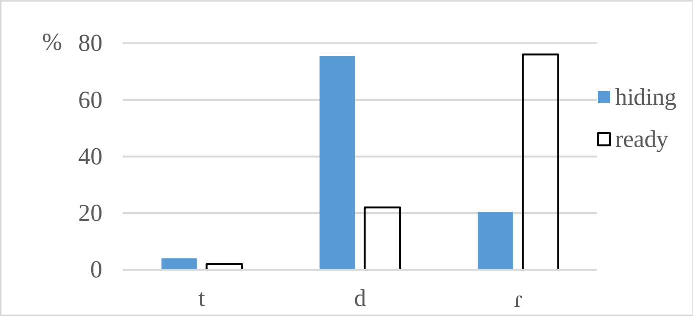
	[t] [d] [ɾ] candidate	Ø	[n]	[ʔ]	others [12]	total
hiding words	76 (58.9)	34 (26.4)	12 (9.3)	1 (0.8)	6 (4.7)	129 (100)
ready words	78 (60)	47 (36.2)	0	0	5 (3.8)	130 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed.

[12] As was the case for /t/ in *eating* and *better* words, [l] accounted for the largest proportion of the variants included in "others" category for *hiding* and *ready* words.

Focusing on how the [t], [d], [ɾ] candidates were actually coded according to the cutoff values, Table 10 presents the frequencies of each variant in *hiding* and *ready* words.

Table 10 The number of tokens judged as [t], [d] and [ɾ] for /d/ in *hiding* and *ready* words in child speech



	[t]	[d]	[ɾ]	total
hiding words	2 (4.1)	37 (75.5)	10 (20.4)	49 (100)
ready words	1 (2)	11 (22)	38 (76)	50 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant.

First, a small number of [t] variants were found in both morphological contexts. This is unexpected, considering that word-medial /d/ is never realized as [t] in the input data, nor is VtV phonetically more natural than VdV. As noted earlier, research suggested that mean percentage of closure in voiced stops evidencing voicing in children's speech is lower than that of adults (Hanson & Shattuck-Hufnagel, 2011; Koenig & Lucero, 2008; B. L.

Smith, 1979). I speculate that the occurrence of [t] in *hiding* and *ready* words is due to such characteristic of child speech.

The probability of producing [d] was much higher in *hiding* words than in *ready* words. As was the case in the analysis of word-medial /t/, it was necessary to consider the age from which the tokens were extracted. Note that the mean age at which *hiding* words were produced was 810 days, while the mean age from which *ready* words were found was 949 days; the difference was statistically significant ($p<0.001$). Clearly, one reason for the higher probability of producing [d] in *hiding* words would be that the articulatory challenges involved in producing flaps has led the children to produce [d] more frequently during the early stage of learning.

In order to examine if OO-faithfulness constraints also contributed to the higher frequency of [d] in *hiding* words, it was necessary to compare the frequencies of each variant in *hiding* and *ready* words within the same learning stage. If OO-faithfulness constraints require that the pronunciation of *hiding* resemble that of *hi*[d]*e*, the dominance of [d] over [ɾ] in *hiding* words must be somehow more robust than in *ready* words, since children's grammar in addition to the articulatory factors would enhance the probability of producing [d].

Figure 5 and Figure 6 show the distribution of [t], [d] and [ɾ] tokens across children's age for /d/ in *hiding* words and *ready* words, respectively.

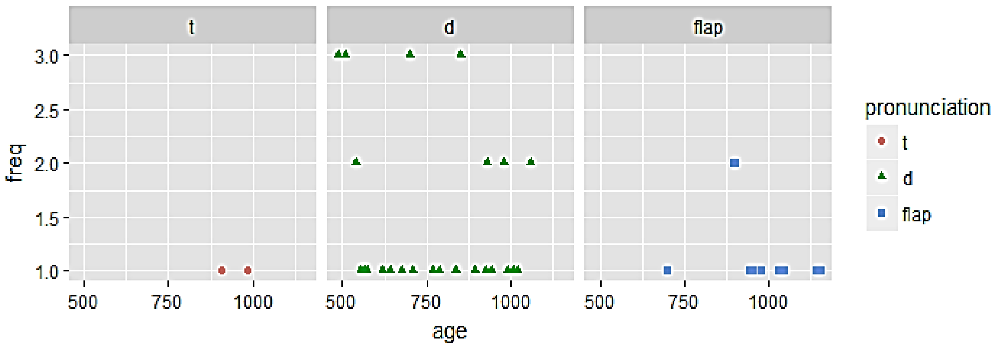


Figure 5 Distribution of [t], [d] and [ɾ] for *hiding* words in children's speech across children's age (in days)

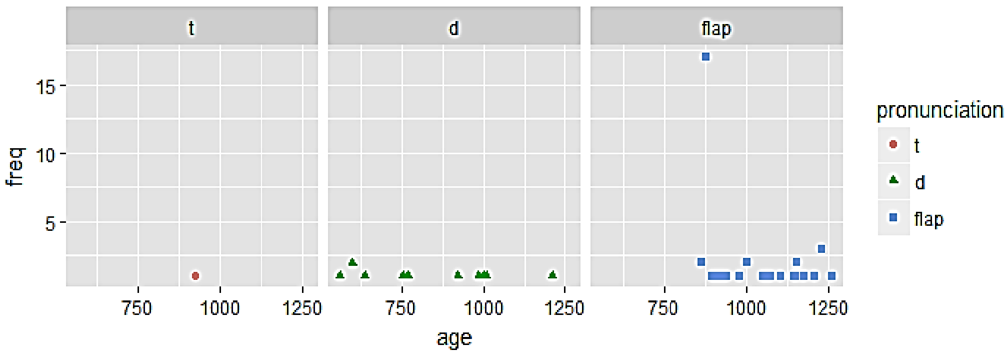


Figure 6 Distribution of [t], [d] and [ɾ] for *ready* words in children's speech across children's age (in days)

In both morphological conditions, flaps were rarely attested in early periods, confirming the findings from previous studies that the probability of producing flaps increases over time as children become experienced learners (Klein & Altman, 2002; Rimac & Smith, 1984). As shown in Table 11, the learning process is divided into three to display the change in frequencies of each variant more clearly.

Table 11 Frequencies of [t], [d] and [ɾ] for *hiding* and *ready* words produced by children in three stages

		Early	Intermediate	Later
hiding words	[t]	0	2 (7.7)	0
	[d]	20 (90.9)	15 (57.7)	2 (40.0)
	[ɾ]	2 (9.1)	9 (34.6)	3 (60.0)
	total	22 (100)	26 (100)	5 (100)
ready words	[t]	0	1 (3.0)	0
	[d]	6 (100)	4 (12.1)	1 (5.6)
	[ɾ]	0	28 (84.8)	17 (94.4)
	total	6 (100)	33 (100)	18 (100)

Note: The numbers in parentheses indicate the percentage of tokens coded as each variant within each stage.

In the Early stage, [d] was predominant and [ɾ] was rarely attested in both *hiding* and *ready* words, reflecting the articulatory challenges that children face in producing flaps. In the Later stage, the frequency of [ɾ] outnumbered that of [d] in both conditions (though there were only a few tokens in *hiding* words), indicating that the children now became to be able to produce flaps actively.

Importantly, in the Intermediate stage, the relative frequencies of each variant are very different in the two morphological contexts. While the frequency of [d] is higher than that of [ɾ] in *hiding* words, the frequency of [d] was lower than that of [ɾ] in *ready* words. Pearson's Chi-squared test with Yates' continuity correction showed that *hiding* and *ready* words significantly differed in their probability of being realized with [d] (χ^2=13.145, df=1, p<0.001). We can notice that the dominance of [d] over [ɾ] in terms of frequency persisted for a longer period in *hiding* words than in *ready* words. Articulatory factors cannot explain why such difference arises, since they are supposed to influence both *hiding* and *ready* words. Therefore, the more prevalent occurrence of [d] in *hiding* words suggests that *hiding* words are additionally regulated by OO-faithfulness constraints that demand for phonological resemblance between a base form (e.g., *hi*[d]*e*) and its suffixed counterpart (e.g., *hi*[d]*ing*).

A question we must ask at this point is whether the base forms such as *hide* are in fact produced with [d], the phonetic variant that is attested in the suffixed forms. Only when /d/ in *hiding* and *hide* words are phonologically similar could we support the hypothesis that children's production of *hi*[d]*ing* is due to OO-faithfulness constraints. I examined the phonetic realizations of word-final /d/ in *hide* words. As was the case for /t/ in *eat* words, only the tokens that appeared in utterance-final position were analyzed.

Table 12 presents attested variants of /d/ in *hide* words. It should be noted that the tokens whose /d/ was realized with an alveolar stop were invariably coded as [d] for convenience, despite some variability in how much voicing was realized during the closure. It has been widely reported in the literature that children often devoice word-final voiced obstruents even in languages like English that do not have automatic final devoicing process (e.g., B. L. Smith, 1979). Given that children lack the ability to maintain vocal-fold vibration and that underlying voiced obstruents are often realized as voiceless (or partially voiced) in child speech, I expected that /d/ for *hide* words in child speech can be phonetically [t]-like.

Table 12 Realization of word-final /d/ in *hide* words produced by children

	[d] [13]	Ø	others [14]	total
hide words	140 (71.4)	46 (23.5)	10 (5.1)	196 (100)

Note: The numbers in parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed.

It was found that the probability of producing word-final /d/ with an alveolar stop was not extremely high, because there were quite a lot of cases (23.5%) where the coda stop was omitted in production. This is not surprising, as it is

[13] This category includes [d] with a clear release (61.2%), [d̚] (6.1%) and [dV] (4.1%), where [V] denotes an epenthetic vowel.
[14] Variants in the "others" category included [ʔ] and some fricatives.

well known in the acquisition literature that children often delete coda consonants (e.g., Fikkert, 1994). The deletion rate was even higher than that of word-final /t/ in *eat* words, which was 10.6 percent (see Table 7). However, when children did produce the coda, the majority of the tokens were produced with an alveolar stop. In this sense, it seems that children were aware that /d/ in *hide* should be realized as [d] and that their production of *hi*[d]*ing* was indeed motivated by its phonological resemblance to the base form *hi*[d]*e*.

Table 13 shows variants of /d/ in *hide* words produced by children in each stage.

Table 13 Realization of word-final /d/ in *hide* words in three stages

	Early	**Intermediate**	**Later**
[d]	67 (76.1)	41 (58.6)	32 (84.2)
Ø	21 (23.9)	21 (30.0)	4 (10.5)
others	0	8 (11.4)	2 (5.3)
total	88 (100)	70 (100)	38 (100)

Note: The numbers in parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed within each stage.

As it was found that the children's tendency to model their pronunciation of /d/ in suffixed forms on that of corresponding base forms persisted until the Intermediate stage, we must check if *hide* words in the first two stages are in fact produced with [d]. Again, when coda stop *is* produced, it is most frequently produced as [d] in all stages.

To summarize, the ratio of [d] to [ɾ] is higher in *hiding* words compared with *ready* words because the probability of producing [d] in suffixed words is enhanced by both articulatory and grammatical factors. We can thus infer from the attested variants of /d/ that children do produce suffixed words that are phonologically similar to their base forms. In the following section, I briefly comment on the assumption in this study that the unsuffixed words serve as the base, rather than the suffixed ones.

4.2.3. Is *Ea*[t] Really the Base?

Since most of the suffixed words investigated in this study are verbs with the progressive inflection, where the suffix -*ing* is attached to verb stems, I primarily focus the discussion on acquisition of verbal paradigm in this section. (See footnote 15 for a brief comment on *bird*-*birdy* pair, which does not belong to this category.) I have assumed so far that among morphologically related forms in a verbal paradigm, the unsuffixed form serves as the base. However, as correctly identifying the intended meaning of children's speech is often very hard, one might well ask if such morphological knowledge is actually present in children's grammar. In this section, I provide some evidence for the validity of this assumption by reviewing previous studies on English-learning children's acquisition of verbal paradigm and by examining the children's speech in the current corpus.

In order to discuss phonological resemblance of a suffixed form to its base in a certain learning stage, one must verify the fundamental assumption of this study that child speakers recognize that base forms and their suffixed counterparts are morphologically related. A number of previous studies have argued that at the very beginning, children do not analyze morphologically complex forms as consisting of more than two elements; in other words, they treat a morphological marker (e.g., -*ing*) and the stem that it attaches to as a single unit (Karmiloff-Smith, 1986; Peters, 1983; Tomasello, 1992). However, these studies rarely provided information on the specific age at which young learners first begin to perform morphological parsing.

In the current data, there is rather weak evidence that children might initially treat suffixed forms as one linguistic chunk. As can be seen in Figure 3, a few tokens of [d] for *eating* words were observed when children were very young. It is possible that children did not recognize, for example, that *eating* consists of the stem *eat* plus the suffix -*ing*, and that they intended to produce *ea*[ɾ]*ing* as adults do, but ended up producing [d]. (That is, without an attempt to model the pronunciation of *eating* on that of *eat*.) If this is the case, there is no point in discussing alternation in children's speech.

However, previous studies provide evidence showing that children soon learn to produce verbs with the suffix -*ing* and that they can construct the morphological relationship between the suffixed form and its unsuffixed counterpart. It is well-known that *-ing* is the first verbal suffix acquired by child learners (Brown, 1973; Gülzow, 2003). Moreover, based on the observation that children initially use the unsuffixed form of the verb for various person and number categories while the *-ing* form is reserved for third person singular usage, Gülzow (2003) argued that young learners regard the unsuffixed forms as more general and basic part of the verbal paradigm. Thus, it seems reasonable to assume that unsuffixed forms are represented as the base form in children's grammar.[15]

Then the question is whether the children in the current data were already able to construct the morphological relationship between the base forms and their suffixed counterparts by the time they *started* to produce *-ing* forms. (Note that in the present study even the tokens from very early periods of children's age, i.e., as early as 462 days, were included in the analysis.) Unfortunately, I could not find reliable evidence from previous studies that suggested child speakers regard unsuffixed forms as the base in those stages of learning, as previous research rarely examined the morphological status of unsuffixed and suffixed forms in early grammar when the *-ing* form first appeared in children's speech. Brown's (1973) study only reported the age at

[15] One exception in the target suffixed words that did not constitute part of a verbal paradigm is *birdy*. Although I did not investigate in detail if *bird* is actually the base form of *birdy*, the fact that 5 out of 6 children produced *bird* in an earlier period than *birdy* hints at the default status of *bird*.

which English verbal suffixes are (nearly) mastered; having investigated three children's speech, he argued that it was as early as 1;10 and as late as 3 years old that the -*ing* suffix consistently appeared in child speech with the correct usage. More extensive research on children's morphological knowledge is necessary.

The findings from the present study can shed some light on children's morphology. The results in section 4.2.1 and 4.2.2 suggest that unsuffixed forms serve as the base, as they are inconsistent with other hypothetical ways of morphological parsing. First, if it is the case that children regard as the base what is typically considered the derived form (e.g., *eating*), assuming that OO-faithfulness constraints are at work, phonetic realization of /t, d/ in the unsuffixed forms would be required to resemble that of /t, d/ in their suffixed counterparts, which is [ɾ] in the majority of cases. It follows that if there are instances where word-final /t/ and /d/ in unsuffixed forms (*eat* and *hide* words) were realized as [ɾ], one can argue that the suffixed forms serve as the base in children's grammar. As can be seen in Table 7 and Table 12, *eat* and *hide* words were never realized with a flap. Therefore, it is unlikely that suffixed forms serve as the base in children's grammar and that the pronunciation of unsuffixed forms must resemble that of suffixed forms. The second alternative possibility is that children do not even possess the notion of base. They might treat a base form and its affixed counterpart equally, as if, for example, *eating* and *eat* have the same morphological status. In that case, children would simply mimic the relative frequencies of the attested variants shown in the input when producing *eating*. In section 4.3, I provide evidence that children do not simply mimic the frequencies of phonetic variants observed in the input.

4.2.4. Attested Learning Trajectory

As mentioned before, I divided the children's learning process into three stages, i.e., the Early, Intermediate and Later stage, focusing mainly on noteworthy changes in the frequencies of [t], [d] and [ɾ] in four different conditions. The frequencies of each variant reported in Table 4 and Table 11 are repeated in (4) and a schematized learning trajectory is presented in (5). In this section, I summarize English-learning children's learning trajectory concerning word-medial /t/ and /d/ in derived and non-derived words.

In the Early stage which covers from 450 to 800 days, [t] was observed for /t/ in suffixed words, suggesting the effect of OO-faithfulness constraints in the initial stage of learning. In this stage, the frequency of [d] was higher than that of [ɾ] regardless of the underlying consonant (i.e., /t/ or /d/) and morphological composition of the target word. In the Intermediate stage, ranging from 800 to 1050 days, the ratio of [d] to [ɾ] differed between suffixed words and non-derived words. More specifically, while the frequency of [d] outnumbered that of [ɾ] in *eating* and *hiding* words, the frequency of [ɾ] outnumbered that of [d] in *better* and *ready* words. It can be inferred that OO-faithfulness constraints are still active in this stage in children's grammar. In the Later stage, when children were aged between 1050 and 1450 days, the frequency of [ɾ] was higher than that of [d] in all four contexts, suggesting that OO-faithfulness constraints are demoted by this time and also that children have learned to produce flaps more actively with a better command of the articulatory systems.

(4) Attested variants and their frequencies in each stage

	Early	**Intermediate**	**Later**
eating			
[t]	16	1	0
[d]	30	23	6
[ɾ]	13	21	12
better			
[t]	1	0	1
[d]	10	10	1
[ɾ]	2	14	14
hiding			
[t]	0	2	0
[d]	20	15	2
[ɾ]	2	9	3
ready			
[t]	0	1	0
[d]	6	4	1
[ɾ]	0	28	17

(5) Attested learning trajectory

	Early (450-800 days)	**Intermediate** (800-1050 days)	**Later** (1050-1450 days)
eating words	[t] attested **[d]** > [ɾ]	**[d]** > [ɾ]	[d] < **[ɾ]**
better words	**[d]** > [ɾ]	[d] < **[ɾ]**	[d] < **[ɾ]**
hiding words	**[d]** > [ɾ]	**[d]** > [ɾ]	[d] < **[ɾ]**
ready words	**[d]** > [ɾ]	[d] < **[ɾ]**	[d] < **[ɾ]**

The observed learning trajectory reported in this section was modeled by means of learning simulation using maxent grammar, the result of which is presented in Chapter 6.

4.3. Adults' Production of Word-medial /t/ and /d/

4.3.1. Realization of Word-medial /t/ in Adult Speech

Mothers' production of medial /t/ and /d/ was investigated to explore the potential influence of input data on children's production of the target words; the observed pattern of production in child speech might simply be a reflection of how mothers speak to children, rather than motivated by initial biases.

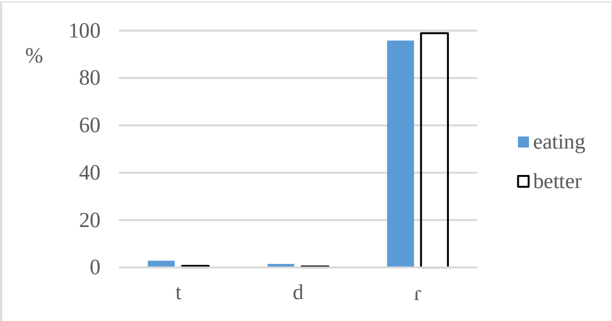
Table 14 shows overall attested variants of /t/ in *eating* and *better* words in mothers' speech and Table 15 shows the number of tokens judged as [t], [d] and [ɾ].

Table 14 Attested variants of /t/ in *eating* and *better* words in adult speech

	[t] [d] [ɾ] candidate	[ʔ]	Ø	others	total
eating words	1041 (81.1)	149 (11.6)	86 (6.7)	8 (0.6)	1284 (100)
better words	564 (95.1)	4 (0.7)	25 (4.2)	0	593 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed.

Table 15 The number of tokens judged as [t], [d] and [ɾ] for /t/ in *eating* and *better* words in adult speech



	[t]	[d]	[ɾ]	total
eating words	20 (2.8)	10 (1.4)	691 (95.8)	721 (100)
better words	3 (0.7)	2 (0.4)	456 (98.9)	461 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant.

Compared to child speech, much less variability is observed, shown in the higher proportions of [t], [d], [ɾ] candidates out of all tokens analyzed (cf. Table 2) and also in the large number of tokens coded as flap (cf. Table 3). Crucially, [t] variants were also attested in mothers' speech, though the frequency was not high. This confirms findings from previous studies, which reported very high rates of flapping in VCV̆ context with occasional cases where /t/ or /d/ surfaces unaltered (Boersma & Hayes, 2001; Hong, 2009). In *eating* words, 20 tokens were realized with [t] accounting for 2.8% of the total number of tokens, and only 3 tokens were realized as such in *better* words, accounting for 0.7% of the total.

It was unexpected that [d] variants were found as realization of /t/ in adult speech. There are a number of possible explanations for this unusual occurrence of [d]. First, as suggested in Song et al. (2015), slower speaking rate of child-directed speech (Fernald & Simon, 1984) might have allowed more time for the occlusion and release of a consonant, giving rise to phonetic [d] variants. Alternatively, it might have been the case that the tokens that were coded as [d] were actually flaps, but were (mis)categorized as [d] according to the cutoff values. Still another possibility is that mothers intentionally mimicked the children's pronunciation of [d].[16]

---

[16] In some cases, it was obvious that mothers intentionally imitated the way children speak. For example, /t/ for *little* in the following utterance was realized with a clear [d], "Naima [Child 4] looked at the picture … and said looking a water little girl", in which Mother 4 was trying to repeat what her daughter had said earlier.

Figure 7 and Figure 8 show the distribution of [t], [d] and [ɾ] tokens in mothers' speech for *eating* and *better* words, respectively.

Figure 7 Distribution of [t], [d] and [ɾ] for *eating* words in mothers' speech across children's age (in days)

Figure 8 Distribution of [t], [d] and [ɾ] for *better* words in mothers' speech across children's age (in days)

It should be noted that [t] variants for *eating* words were found only in the early periods. Assuming that mothers' production of [t] in *eating* words cannot be due to undominated OO-faithfulness constraints, it is possible that mothers produced [t] in *eating* words in order to inform the children of the fact that base forms and their suffixed counterparts are morphologically and semantically related. Meanwhile, mothers' production of [t] in those periods could be an artifact of slower speaking rate in child-directed speech (CDS). Provided that speaking rate in CDS gradually increases starting from the early periods of children's speech production up to a certain developmental stage (Ko, 2012), a particularly slower speaking rate in the beginning stage of speech production might have led to a higher probability of producing [t]. I do not investigate in detail whether mothers actually produced [t] with a higher frequency in their CDS compared with adult-directed speech and what motivated their production of [t] in their speech directed to children in early stage of learning.[17]

More importantly, we should ask, at this point, whether children's production of [t] was an attempt to mimic the way mothers speak. It was found that both children and mothers produced [t] variants and that the periods in which they did so coincided, i.e., early in the learning process. This finding suggests the possibility that mothers produced [t] for *eating* words only when the children were less experienced in learning, and during that period children also produced [t] in order to mimic the frequencies of each variant in mothers'

[17] The percentage of tokens that were coded as [t] in mothers' speech was 2.8% in this study. This does not seem to differ much than the proportion of [t] variants found in other corpora, i.e., 2.84% in TIMIT and 2.51% in Buckeye Speech Corpus, reported in Hong (2009). For now, therefore, there is no concrete evidence for a higher probability of producing [t] in child-directed speech compared with adult-directed speech. (Note that a direct comparison of the proportion of [t] between the present study and Hong (2009) is possible, since both studies calculated the proportion in the same way, i.e., the number of tokens coded as [t] divided by the number of tokens coded as [t] or [ɾ] (and also [d] in the current study), not by the total number of tokens observed in each corpus.)

speech. Under this scenario, we would have no evidence for the claim that children produced variants that are faithful to the base form due to high-ranked OO-faithfulness constraints. Note, however, that the probability of producing [t] was higher in the children's speech (13.9%, see Table 3) than in the mothers' speech (2.8%). Therefore, it does not seem to be the case that children simply mimicked the relative frequencies of each variant observed in the input.

In order to investigate the possible influence of the input more accurately, I compared the probability of producing [t] in the speech of Child 4 and Mother 4. As most of the [t] tokens were found in the speech of Child 4, a more precise comparison must concern only Child 4 – Mother 4 dyad. As it turns out, most of the [t] tokens attested in mothers' speech actually came from the speech of Mother 4. Table 16 presents frequencies of [t], [d] and [ɾ] in *eating* words produced by each child and mother.

Table 16 Frequencies of [t], [d] and [ɾ] for *eating* words produced by each child and mother

		MOT 1	MOT 2	MOT 3	MOT 4	MOT 5	MOT 6
eating words	[t]	0	0	1	18	0	1
	[d]	0	1	4	5	0	0
	[ɾ]	23	77	181	323	16	71
		CHI 1	CHI 2	CHI 3	CHI 4	CHI 5	CHI 6
	[t]	0	0	1	16	0	0
	[d]	5	5	11	36	2	0
	[ɾ]	0	0	6	40	0	0

It was found that while Child 4 produced [t] in 16 out of 92 tokens (17.4%), Mother 4 did so in 18 out of 346 tokens (5.2%). A Pearson's Chi-squared test with Yates' continuity correction showed that the difference is statistically significant (χ^2=13.426, df=1, p<0.001). It was also revealed that the frequency of [t] produced by Child 4 in the Early stage (15 out of 50) was significantly different from that of Mother 4 in the same period (18 out of 265) (χ^2=21.745, df=1, p<0.001). Therefore, I conclude that although there are reasons to believe that mothers' speech affected the children's production to some extent, the children did not directly replicated the relative frequencies of phonetic variants observed in the input.

4.3.2. Realization of Word-medial /d/ in Adult Speech

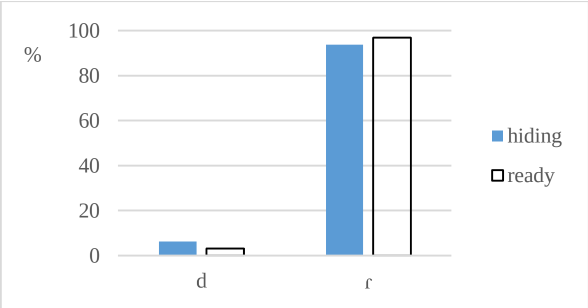
Mothers' production of /d/ in *hiding* and *ready* words was also analyzed. It is shown in Table 17 that over 90% of the total number of tokens could be coded as [t], [d] or [ɾ] in both contexts. Table 18 shows that among these tokens most were coded as [ɾ], suggesting a high rate of flapping as expected. It was also found that both *hiding* and *ready* words were realized with [d] in a small number of tokens. The probability of producing [d] was higher in *hiding* words (6.3%) than in *ready* words (3.1%). Mothers might have produced [d] more frequently for *hiding* words in order to provide the children with phonological cues to morphological relationship between a suffixed word and its corresponding base form, e.g., *hiding* and *hide.* This explanation seems more likely when we consider the age at which [d] was produced in *hiding* words.

Table 17 Attested variants of /d/ in *hiding* and *ready* words in adult speech

	[t] [d] [ɾ] candidate	Ø	others	total
hiding words	346 (92.3)	27 (7.2)	2 (0.5)	375 (100)
ready words	658 (91.6)	55 (7.7)	5 (0.7)	718 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant out of all tokens analyzed.

Table 18 The number of tokens judged as [t], [d] and [ɾ] for /d/ in *hiding* and *ready* words in adult speech



	[t]	[d]	[ɾ]	total
hiding words	0	13 (6.3)	194 (93.7)	207 (100)
ready words	0	13 (3.1)	403 (96.9)	416 (100)

Note: The numbers in the parentheses indicate the percentage of tokens coded as each variant.

Figure 9 and Figure 10 show the distribution of [d] and [ɾ] variants across children's age for *hiding* and *ready* words, respectively. It is shown in Figure 9 that [d] for *hiding* words was mostly attested in the early stage of learning. (Recall that it was also the case in *eating* words that the variant that is faithful to the base, i.e., [t], was found only in the early periods in mothers' speech.) Thus, one might speculate that such phonological similarity between morphologically related forms reflects mothers' tendency to speak more clearly when children were younger. However, a Pearson's Chi-squared test with Yates' continuity correction revealed that the probability of producing [d] in *hiding* words did not significantly differ from that of *ready* words (χ^2=2.697, df=1, p=0.1005), indicating that the frequency of [d] was not particularly high in derived words.

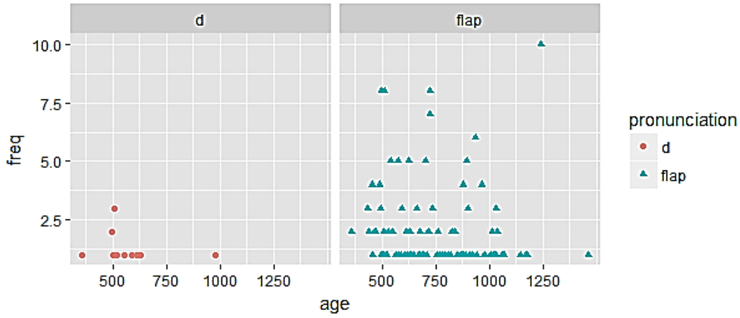


Figure 9 Distribution of [d] and [ɾ] for *hiding* words in mothers' speech across children's age (in days)

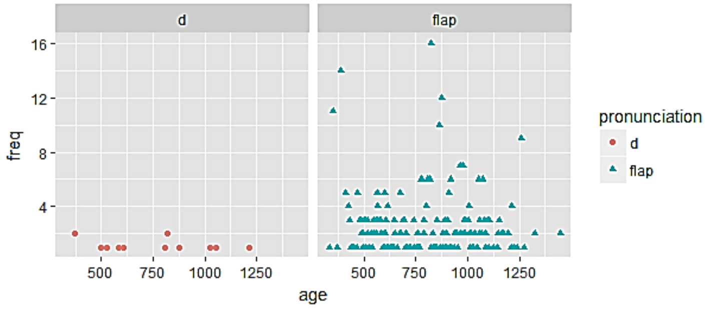


Figure 10 Distribution of [d] and [ɾ] for *ready* words in mothers' speech across children's age (in days)

We now turn to the question of whether or not children's frequent production of [d] in *hiding* words was influenced by the input. A more careful comparison between children's and mothers' speech indicated that it is unlikely that children mimicked the frequencies of the variants attested in their mothers' speech. First, *hiding* and *ready* words significantly differed in their probability of being realized with [d] only in children's speech, but not in mothers' speech. Another piece of evidence against the significant influence of input on children's production of [d] in *hiding* words is concerned with the learning stage in which children produced *hi*[d]*ing*. More specifically, if we assume that the children's probability of producing each variant was determined by the relative frequencies observed in the input, there is no way to explain the dominance of [d] over [ɾ] in *hiding* words in the Intermediate stage. In other words, as mothers produced *hi*[d]*ing* mostly in the Early stage but children continued to produce such variant in the Intermediate stage, it is hard to claim that children’s production of *hi*[d]*ing* is primarily influenced by the input. I conclude that children’s tendency to produce non-alternating forms is mainly due to the OO-faithfulness constraints, though the influence of input might have contributed to the tendency at least partially.

In sum, the higher frequency of [d] for *hiding* words in child speech cannot be explained by the influence of input, suggesting that it is instead motivated by OO-faithfulness constraints. The results for /t/ and /d/ taken together, it was found that the observed frequencies of each variant for /t/ and /d/ in the children’s speech do not closely match those found in the mothers’ production. Therefore, I conclude that children’s tendency to produce suffixed words with phonetic variants that are faithful to the corresponding segment in the base form, especially in earlier stages of learning, is due to high-ranked OO-faithfulness constraints at the initial state of learning and its demotion in the learning process.

5. Analysis: An Optimality Theoretic Account

In this chapter, OT constraints are introduced that can explain how /t/ and /d/ are realized in child speech. I only consider candidates that can be considered typical variants of /t/ and /d/, i.e., [t], [d] and [ɾ], which accounted for a significant portion of the surface realizations of the stops in both children or mothers' speech.

5.1. Flapping

First, constraints are needed that trigger flapping. It is well known in the literature that flap is distinguished from the alveolar stops [t] and [d] by its extra-short closure duration (Banner-Inouye, 1995; Steriade, 2000; Zue & Laferriere, 1979). I assume that the durational difference between flap and the stops can explain why [t] and [d] are banned but a flap is allowed in certain phonological contexts. I posit two markedness constraints that trigger flapping. The one in (6) militates against intervocalic [t] and the one in (7) militates against intervocalic [d] that is followed by an unstressed vowel.

(6) *VtV̌

No intervocalic alveolar voiceless stops followed by an unstressed vowel.

(7) *VdV̌

No intervocalic alveolar voiced stops followed by an unstressed vowel.

The rationale for establishing two distinct constraints concerned with flapping, one for /t/ and the other for /d/, is first because intervocalic voiceless consonants can be considered more marked than voiced consonants in the same environment. In addition, it was confirmed in the present study that children in general favored [d] to [t] in VCV̌ context, shown in the high probabilities of producing [d] in all of the four environments, i.e., /t/ and /d/ in derived and non-derived words. It is thus expected that (6) is assigned a higher weight than (7) in learning simulation.

Considering that flaps only appear in restricted environments in English, we need a context-free markedness constraint that bans flaps. The constraints in (8) penalize candidates that contain [ɾ].

(8) *ɾ

No flaps.

When underlying /t/ or /d/ surfaces as a flap, IDENT constraints are violated. As noted in Bernhardt and Stemberger (1998), the relevant constraints are ID-IO(son) and ID-IO(voi). When /t/ surfaces as a flap, both constraints will be violated (assuming that flaps are [+sonorant], [+voice]); when /d/ surfaces as a flap, only ID-IO(son) will be violated.

(9) ID-IO(son)

The value of the feature [sonorant] of an input segment must be preserved in its output correspondent.

(10) ID-IO(voi)

The value of the feature [voice] of an input segment must be preserved in its output correspondent.

The constraints introduced so far can explain the general phonology in American English, the flapping process in particular.[18] The tableaux in (11)

---

[18] In addition to the two IDENT constraints for [voice] and [sonorant], I could also introduce an IDENT constraint for the feature [extra-short closure]. Since I assumed that the motivation for the constraints in (6) and (7) is the durational difference between flap and the stops, it seems appropriate to establish IDENT constraints for the durational feature. Steriade (2000) also argued that it is not satisfactory to use [sonorant] in distinguishing flap from [t]/[d] when we have to capture the connection between occurrence of a flap and the shortening context that triggers it. However, I will only consider phonologically contrastive features [voice] and [sonorant], following Hayes (2004) and Bernhardt and Stemberger (1998). Even if I did include IDENT constraints for [extra-short closure], its effect on constraint evaluation would not be very critical, since it plays exactly the same role as ID-IO(son).

and (12) illustrate a set of analyses based on the constraints in (6)-(10) with *eating* and *hiding*. As will be discussed in Chapter 6, maxent grammar, the theoretical framework used in this thesis, assigns numerical weights to each constraint and probability distributions over variants. Therefore, in the tableaux below, no single candidate is designated as a winner, nor is the constraint ranking fixed (as represented by dotted lines).

(11)

/iːtɪŋ/	*VtV̆	*VdV̆	*ɾ	ID-IO(son)	ID-IO(voi)
iːtɪŋ	*				
iːdɪŋ		*			*
iːɾɪŋ			*	*	*

(12)

/haɪdɪŋ/	*VtV̆	*VdV̆	*ɾ	ID-IO(son)	ID-IO(voi)
haɪ.tɪŋ	*				*
haɪ.dɪŋ		*			
haɪ.ɾɪŋ			*	*	

5.2. OO-faithfulness Constraints

OO-faithfulness constraints must be established in order to explain the children's tendency to base their pronunciation of derived words (e.g., *eating*) on the corresponding base forms (e.g., *eat*). Analogous to the constraints in

(9) and (10), two OO-faithfulness constraints are introduced, each targeting [sonorant] and [voice]. The constraints in (13) and (14) were assumed in Bernhardt and Stemberger (1998) and Hayes (2004) in explaining the young learner's bias for non-alternation concerning flapping.

(13) ID-OO(son)

The value of the feature [sonorant] of a segment in the base must be preserved in its correspondent in the affixed form.

(14) ID-OO(voi)

The value of the feature [voice] of a segment in the base must be preserved in its correspondent in the affixed form.

As (13) and (14) require phonological similarity among output forms, they regulate suffixed-unsuffixed word pairs such as *eating* and *eat*, but they have no effect on words such as *better* which have no base form. The tableaux in (15)-(17) illustrate analyses based on constraints adopted so far with *eating*, *eat* and *better*.

(15)

/iːtɪŋ/ base: [iːt]	*VtV̆	*VdV̆	*ɾ	ID-OO (son)	ID-OO (voi)	ID-IO (son)	ID-IO (voi)
iːtɪŋ	*						
iːdɪŋ		*			*		*
iːɾɪŋ			*	*	*	*	*

(16)

/iːt/	*VtV̆	*VdV̆	*ɾ	ID-OO (son)	ID-OO (voi)	ID-IO (son)	ID-IO (voi)
iːt							
iːd							*
iːɾ			*			*	*

(17)

/betəɹ/ base: none	*VtV̆	*VdV̆	*ɾ	ID-OO (son)	ID-OO (voi)	ID-IO (son)	ID-IO (voi)
be.təɹ	*						
be.dəɹ		*					*
be.ɾəɹ			*			*	*

Note that the flap candidate in (15) violates OO-faithfulness constraints, while the flap candidates in (16) and (17) vacuously satisfy them. Thus, the effect of OO-faithfulness constraints is to discourage the probability of producing flaps for /t/ in derived words, compared with that of non-derived words. Similarly, the tableaux in (18)-(20) illustrate analyses with *hiding*, *hide* and *ready*.

(18)

/haɪdɪŋ/ base: [haɪd]	*VtV̆	*VdV̆	*ɾ	ID-OO (son)	ID-OO (voi)	ID-IO (son)	ID-IO (voi)
haɪ.tɪŋ	*				*		*
haɪ.dɪŋ		*					
haɪ.ɾɪŋ			*	*		*	

(19)

/haɪd/	*VtV̆	*VdV̆	*ɾ	ID-OO (son)	ID-OO (voi)	ID-IO (son)	ID-IO (voi)
haɪt							*
haɪd							
haɪɾ			*			*	

(20)

/redi/ base: none	*VtV̆	*VdV̆	*ɾ	ID-OO (son)	ID-OO (voi)	ID-IO (son)	ID-IO (voi)
re.ti	*						*
re.di		*					
re.ɾi			*			*	

The constraints established so far are needed in explaining the children's production of /t/ and /d/, and were used in learning simulation to learn the children's grammar in each stage. Table 19 summarizes the set of constraints adopted to analyze the children's production of word-medial /t/ and /d/.

Table 19 Constraint set adopted for the analysis of realization of /t/ and /d/ in child speech

Constraint	*Interpretation*
a. *VtV̆	No intervocalic alveolar voiceless stops followed by an unstressed vowel.
b. *VdV̆	No intervocalic alveolar voiced stops followed by an unstressed vowel.
c. *ɾ	No flaps.
e. ID-IO(son)	The value of the feature [sonorant] of an input segment must be preserved in its output correspondent.
f. ID-IO(voi)	The value of the feature [voice] of an input segment must be preserved in its output correspondent.
g. ID-OO(son)	The value of the feature [sonorant] of a segment in the base must be preserved in its correspondent in the affixed form.
h. ID-OO(voi)	The value of the feature [voice] of a segment in the base must be preserved in its correspondent in the affixed form.

6. Learning Simulation

Based on the constraints adopted in the previous chapter, a set of learning simulation was conducted using maxent grammar tool (Hayes, 2009) in order to learn the children's grammar in each stage. In maxent model (Goldwater & Johnson, 2003), one version of probabilistic OT, constraint weights are determined in a way that maximizes the probability of the observed frequencies of output forms. Unlike in standard OT where a fixed constraint ranking is assumed, maxent grammar assigns numerical weights to each constraint. More specifically, constraints that are violated by frequently observed output forms are weighted with a lower value, while those violated by less frequent output forms are weighted higher. With the constraint weights thus learned, the learner predicts and assigns probabilities to each output. When provided with an adequate constraint set, the grammar is able to reproduce predicted frequencies of each candidate that closely match observed frequencies of each form. Maxent grammar has an advantage in accounting for variation, in that it uses weighted constraints to assign probabilities to outputs, rather than choosing one candidate as an optimal output.

Three independent learning simulations were conducted, each of them learning the children's grammar in the Early, Intermediate, and Later stage.

Learning data, shown in Table 20, were constructed based on the children's production data for /t/ and /d/ in derived words (e.g., *eating*), their base forms (e.g., *eat*) and non-derived words that were not part of a paradigm (e.g., *better*). The learner was run in a default setting (mu=0, sigma^2=100,000 for all constraints). The value of mu indicates the prior expected mean weight of a constraint. Sigma value is the variance of the weight; lower sigma value means being more conservative about constraint weights.

Table 20 Learning data for children's production of /t/ and /d/

Input	Output	Early	Inter.	Later	*VtV̆	*VdV̆	*ɾ	ID-IO (son)	ID-IO (voi)	ID-OO (son)	ID-OO (voi)
eating	[t]	16	1	0	*						
	[d]	30	23	6		*			*		*
	[ɾ]	13	21	12			*	*	*	*	*
eat	[t]	124	46	25							
	[d]	0	0	0					*		
	[ɾ]	0	0	0			*	*	*		
better	[t]	1	0	1	*						
	[d]	10	10	1		*			*		
	[ɾ]	2	14	14			*	*	*		
hiding	[t]	0	2	0	*				*		*
	[d]	20	15	2		*					
	[ɾ]	2	9	3			*	*		*	
hide	[t]	0	0	0					*		
	[d]	67	41	32							
	[ɾ]	0	0	0			*	*			
ready	[t]	0	1	0	*				*		
	[d]	6	4	1		*					
	[ɾ]	0	28	17			*	*			

Table 21 presents the constraint weights learned. In each stage, the constraints are presented in descending order of the value of weights.

Table 21 Constraint weights learned

Early		**Intermediate**		**Later**	
*VtV̆	24.684	*VtV̆	16.692	*VtV̆	26.070
ID-IO(voi)	12.458	*VdV̆	12.092	*VdV̆	12.898
*VdV̆	9.971	*ɾ	5.503	ID-IO(voi)	11.243
*ɾ	5.664	ID-IO(son)	5.503	*ɾ	5.249
ID-IO(son)	5.664	ID-IO(voi)	2.609	ID-IO(son)	5.249
ID-OO(voi)	1.496	ID-OO(son)	1.316	ID-OO(son)	1.666
ID-OO(son)	0.000	ID-OO(voi)	0.000	ID-OO(voi)	0.000

In all stages, *VtV̆ was assigned the largest weight among all constraints, accounting for the fact that children rarely produced [t] in intervocalic position throughout the learning process. Even though the probability of producing a flap for the four types of target words, i.e., *eating*, *better*, *hiding* and *ready* words, increased along the learning process, *ɾ was assigned a relatively high weight even in the Later stage. This seems to be because the base forms, i.e., *eat* and *hide* words, never surfaced with a flap. As the frequency of the output forms (or the number of tokens found in the corpus) for the base forms was in general higher than that of the four types of target words (see Table 20), their effect on the constraint weight might have been substantial, leading to a relatively high weight of *ɾ.

It was found that in all stages the two OO-faithfulness constraints were assigned a low weight or zero. First, ID-OO(voi) had a weight of 1.496 in the Early stage, and it was assigned zero weight in the Intermediate and the Later stage. As ID-OO(voi) penalizes [d] and [ɾ] as realizations of /t/ in suffixed forms, the fact that it was assigned some weight only in the Early stage explains that children's production of [t] in the derived words (but not in non-derived words) was observed only in this stage. (But the frequency of [t] was not extremely high even in the Early stage, as reflected in the large weight of *VtV̆.) As ID-OO(voi) that bans [t]~[d] and [t]~[ɾ] alternation is assigned zero weight in the two latter stages, [d] and [ɾ] frequently occurs in *eating* words. Meanwhile, ID-OO(son) was assigned zero weight in the Early stage, though it was assigned some weight in the Intermediate and the Later stage. This is because the role of ID-OO(son) is to ban flaps in words that had a base form, but in the actual data the probability of producing a flap was discouraged even in words without a base, since flaps are hard to pronounce in this stage, no matter the morphological status of a word. Only when the children gained ability to produce flaps actively (i.e., in the Intermediate and the Later stage) did they exhibit a production pattern in which [d]~[ɾ] alternation was banned in suffixed words, reflected in the weight of ID-OO(son).

The learned grammar was able to reproduce the relative frequencies of variants for /t/ and /d/ in children's production. Table 22 shows the observed and predicted frequencies of phonetic variants of /t/ and /d/ in the four contexts.

Table 22 Observed and predicted frequencies of variants of /t/ and /d/

		Early		**Intermediate**		**Later**	
		observed	predicted	observed	predicted	observed	predicted
eating	[t]	0.271	0.271	0.022	0.071	0.000	0.045
	[d]	0.508	0.580	0.511	0.518	0.333	0.309
	[ɾ]	0.220	0.149	0.467	0.411	0.667	0.646
eat	[t]	1.000	1.000	1.000	0.931	1.000	1.000
	[d]	0.000	0.000	0.000	0.069	0.000	0.000
	[ɾ]	0.000	0.000	0.000	0.000	0.000	0.000
better	[t]	0.077	0.077	0.000	0.033	0.063	0.012
	[d]	0.769	0.734	0.417	0.244	0.063	0.082
	[ɾ]	0.154	0.189	0.583	0.722	0.875	0.906
hiding	[t]	0.000	0.000	0.077	0.000	0.000	0.000
	[d]	0.909	0.795	0.577	0.557	0.400	0.324
	[ɾ]	0.091	0.205	0.346	0.442	0.600	0.676
hide	[t]	0.000	0.000	0.000	0.069	0.000	0.000
	[d]	1.000	1.000	1.000	0.931	1.000	1.000
	[ɾ]	0.000	0.000	0.000	0.000	0.000	0.000
ready	[t]	0.000	0.000	0.030	0.000	0.000	0.000
	[d]	1.000	0.795	0.121	0.253	0.056	0.083
	[ɾ]	0.000	0.205	0.848	0.747	0.944	0.917

A linear regression analysis with the observed frequency as the dependent variable and the predicted frequency as the independent variable showed that the predicted frequencies can account for 94% of the variance found in the observed frequencies (R^2=0.94, p<0.001). The predictability varies somewhat depending on the stage (Early: R^2=0.93, p<0.001, Intermediate: R^2=0.86, p<0.001, Later: R^2=0.99, p<0.001), but overall the observed and predicted frequencies matched closely, suggesting that the proposed constraint set can adequately explain the patterns found in children's production. Figure 11 demonstrates correlation between the observed and predicted frequencies in the three learning stages.

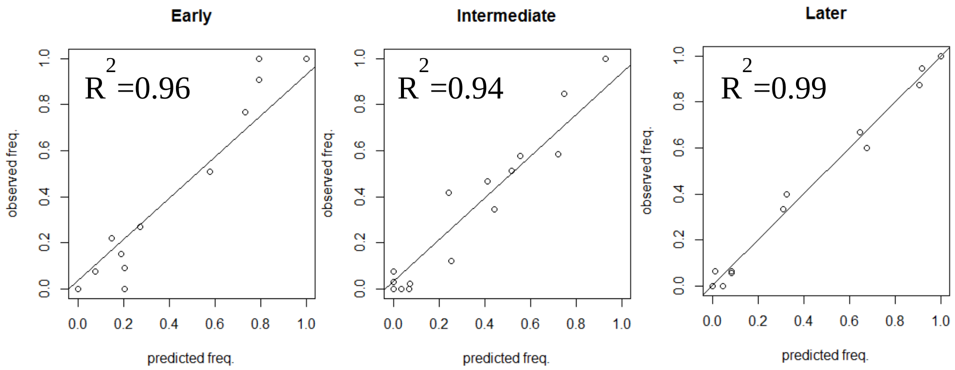


Figure 11 Correlation between the observed frequencies and the predicted frequencies

7. Conclusion

The present study aimed to test the claim in McCarthy (1998) and Hayes (2004) that OO-faithfulness constraints are undominated in the initial stage of learning, but are gradually demoted in the learning process as the young learners encounter violations of those constraints in the input data. With the case of American English Flapping, the production data of six English learning children provided evidence that they indeed showed initial bias for non-alternation, enforced by high-ranked OO-faithfulness constraints, and that such bias became weaker over time. Crucially, this study also examined the influence of input data that children receive from their mothers, a factor that has been neglected in previous studies, thus testing the hypothesis in McCarthy (1998) and Hayes (2004) more rigorously. The investigation of the mothers' speech showed that children's production of variants that are faithful to the base is not a direct replication of the observed frequencies in the input. The attested learning trajectory was modeled with maxent grammar based on OT constraints.

References

Aksu-Koç, A., & Slobin, D. I. (1985). The acquisition of Turkish. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition, vol. 1: the data* (pp. 839–878). Hillsdale, NJ: Erlbaum.

Banner-Inouye, S. (1995). *Trills, taps and stops in contrast and variation* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.

Benua, L. (1997). *Transderivational identity: Phonological relations between words* (Doctoral dissertation, University of Massachusetts, Amherst). Retrieved from https://rucore.libraries.rutgers.edu/rutgers-lib/38308/

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150–177.

Berman, R. A. (1985). The acquisition of Hebrew. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition, vol. 1: the data* (pp. 255–371). Hillsdale, NJ: Erlbaum.

Bernhardt, B. H., & Stemberger, J. P. (1998). *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. San Diego: Academic Press.

Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, *32*(1), 45–86.

Brown, R. (1973). *A first language*. Cambridge: Cambridge University Press.

Byrd, D. (1993). *54,000 American Stops* (UCLA Working Papers in Phonetics No.83). Retrieved from ResearchGate website: https://www.researchgate.net/publication/2460697_54000_American_Stops

Clahsen, H., Aveledo, F., & Roca, I. (2002). The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language*, *29*, 591–622.

Clahsen, H., Prüfert, P., Eisenbeiβ, S., & Cholin, J. (2002). Strong stems in the German mental lexicon: Evidence from child language acquisition and adult processing. In I. Kaufmann & B. Stiebels (Eds.), *More than words. Festschrift for Dieter Wunderlich* (pp. 91–112). Berlin: Akadamie Verlag.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech*, *49*(2), 137–173.

Do, Y. A. (2013). *Biased learning of phonological alternations* (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from https://yd79.files.wordpress.com/2009/09/do_dissertation_mit.pdf

Eddington, D. (2007). Flaps and other variants of /t/ in American English: Allophonic distribution without constraints, rules, or abstractions. *Cognitive Linguisitcs*, *18*(1), 23–46.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20,* 104–113.

Fikkert, P. (1994). *On the acquisition of prosodic structure* (Doctoral dissertation). Retrieved from http://repository.ubn.ru.nl/

Flege, J. E., & Brown, W. S. J. (1982). The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance. *Journal of Phonetics, 10,* 335–345.

Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In J. Spenador, A. Eriksson, & O. Dahl (Eds.), *Proceedings of the workshop on variation within Optimality Theory* (pp. 111–120). Stockholm University.

Gülzow, I. (2003). Early development of verbal morphology in an English-speaking child. In D. Bittner, W. U. Dressler, & M. Kilani-Schoch (Eds.), *Development of verb inflection in first language acquisition: A cross-linguistic perspective* (pp. 205–238). Berlin: De Gruyter Mouton.

Hanson, H. M., & Shattuck-Hufnagel, S. (2011). An acoustic study of word-medial stop consonants in 2-3 year-old speakers of American English. In W-S. Lee & E. Zee (Eds), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)* (pp. 819–822). Hong Kong.

Hayes, B. (2004). Phonological acquisition in Optimality Theory: The early stages. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Fixing priorities: constraints in phonological acquisition* (pp. 158–203). Cambridge: Cambridge University Press.

Hayes, B. (2009). Maxent grammar tool, software.

Hong, S.-H. (2009). Variation and frequency in word-medial /t/-flapping in American English. *Studies in Phonetics, Phonology and Morphology*, *15*(2), 375–396.

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. The Hague: Mouton.

Kahn, D. (1976). *Syllable-based generalizations in English phonology* (Doctoral dissertation). Retrieved from https://dspace.mit.edu

Karmiloff-Smith, A. (1986). Stage/structure versus phase/process in modeling linguistic and cognitive development. In I. Levin (Ed.), *Stage and structure: Reopening the debate* (pp. 164–190). Norwood, NJ: Ablex.

Kazazis, K. (1969). Possible evidence for (near-)underlying forms in the speech of a child. *Papers of the Annual Regional Meeting, Chicago Linguisitc Society*, *5*, 382–388.

Keating, P. A. (1984). Phonetic and Phonological Representation of Stop Consonant Voicing. *Language*, *60*(2), 286–319.

Kenstowicz, M. (1997). Base identity and uniform exponence: Alternatives to cyclicity. In J. Durand & B. Laks (Eds.), *Current trends in* *phonology: Models and methods* (pp. 363–394). Salford: University of Salford.

Kewley-Port, D., & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, *2*, 195–210.

Klein, H. B., & Altman, E. K. (2002). The acquisition of medial /t, d/ allophones in bisyllabic contexts. *Clinical Linguistics & Phonetics*, *16*(3), 215–232.

Klein, H. B., Altman, E. K., & Tate, J. (1998). Relationship between perceptual and acoustic judgments of children's flap productions. *Contemporary Issues in Communication Sciences and Disorders*, *25*, 25–31.

Ko, E. S. (2012). Nonlinear development of speaking rate in child-directed speech. *Lingua*, *122*(8), 841–857.

Koenig, L. L., & Lucero, J. C. (2008). Stop consonant voicing and intraoral pressure contours in women and children. *Journal of Acoustical Society of America*, *123*(2), 1077–1088.

Kreidler, C. W. (1989). *The pronunciation of English: A course book in phonology*. Cambridge, MA: Blackwell.

Macken, M., & Barton, D. A. (1980). Longitudinal study of the acquisition of the voicing contrast in American English word-initial stops, as measured by voice onset time. *Journal of Child Language*, *7*, 41–74.

Magen, H. S., & Blumstein, S. E. (1993). Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics*, *21*, 387–409.

McCarthy, J. J. (1998). Morpheme structure constraints and paradigm occultation. *Papers of the Annual Regional Meeting, Chicago Linguisitc Society*, *32*, 123–150.

Patterson, D., & Connine, C. M. (2001). Variant frequency in flap production. *Phonetica*, *58*, 254–275.

Peters, A. (1983). *The units of language acquisition*. New York: Cambridge University Press.

Prince, A., & Smolensky, P. (1993). *Optimality theory: constraint interaction in generative grammar*. New Brunswick, NJ: Rutgers Center for Cognitive Science, Rutgers University.

Rimac, R., & Smith, B. L. (1984). Acoustic characteristics of flap productions by American English-speaking children and adults. *Journal of Phonetics*, *12*(4), 387–396.

Smith, B. L. (1979). A phonetic analysis of consonantal devoicing in children's speech. *Journal of Child Language*, *6*, 19–28.

Smith, N. (1973). *The acquisition of phonology*. Cambridge: Cambridge University Press.

Song, J. Y., Shattuck-Hufnagel, S., & Demuth, K. (2015). Development of phonetic variants (allophones) in 2-year-olds learning American English: A study of alveolar stop /t, d/ codas. *Journal of Phonetics*, *52*, 152–169.

Steriade, D. (2000). Paradigm Uniformity and the Phonetics-Phonology Boundary. In J. Pierrehumbert & M. Broe (Eds.), *Papers in Laboratory Phonology vol.6* (pp. 313–334). Cambridge: Cambridge University Press.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge, UK: Cambridge University Press.

Wardrip-Fruin, C. (1982). On the status of temporal cues to phonetic categories: Preceding vowel duration as a cue to voicing in final stops. *Journal of the Acoustical Society of America*, *71*, 187–202.

White, J. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, *130*, 96–115.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, *30*, 945–982.

Yang, S. I. (2011). Hankwuke cwungchep piumuy kiliey tayhan kochal [A study on the duration of Korean geminate nasals]. *Korean Linguistics*, *51*, 93–116.

Yao, Y. (2007). Closure duration and VOT of word-initial voiceless plosives in English in spontaneous connected speech. *UC Berkeley Phonology Lab Annual Report*, 183–225.

Zue, V., & Laferriere, M. (1979). Acoustic study of medial /t,d/ in American

English. *Journal of the Acoustical Society of America*, *66*, 1039–1050.

국문초록

영어를 습득하는 아동의 음운 교체에 대한 학습 편향

본 논문은 출력형 대 출력형 충실성 제약(output-to-output faithfulness constraints)이 선험적으로 상위에 위치해 있다는 McCarthy(1998)와 Hayes(2004)의 주장을 검증하고자 한다. 이 주장에 따르면, 출력형 대 출력형 충실성 제약은 학습 초기에 최상위에 위치해 있다가 어린 아이들이 음운 교체(phonological alternation) 현상에 점차 노출됨에 따라 점진적으로 강등된다. 이러한 주장이 사실이라면, 아이들이 학습 초기에는 교체 현상을 회피하는 경향을 보이지만 시간이 지나면 어른 화자와 같이 교체형을 활발하게 발화할 것으로 예상해볼 수 있다.

이와 관련하여, 미국 영어의 어른 화자들은 특정 음운 환경에서 /t/와 /d/를 탄설음 [ɾ]으로 발음하는 반면, 어린 아이들은 패러다임의 평준화(paradigm leveling)를 만족시키기 위해 같은 환경에서 탄설음화(flapping) 규칙을 적용시키지 않는다는 것이 Bernhardt & Stemberger(1998)에서 보고되었다. 예컨대, 아이들은 'sitting'을

si[ɾ]ing 이 아닌 si[t]ing 으로 발음하는데, 이것은 탄설음화를 적용시켜 si[ɾ]ing 이라고 발화할 경우 기본형 si[t]과 파생형 si[ɾ]ing 사이에 [t]~[ɾ] 음운 교체가 발생하기 때문이다. 선행연구의 이러한 관찰에 기반하여 본 연구는 어린 아이들의 발화에서 교체형이 나타나는 비율이 나이에 따라 어떻게 변화하는지 조사한다. 특히, 본 연구는 어머니의 언어 입력(input)이 아이들의 교체 회피 경향에 영향을 줄 가능성을 검토한다. 미국 영어에서 탄설음화 환경에 있는 /t/가 [ɾ]로 실현되지 않고 [t]로 실현되는 경우도 있다는 점을 고려하면, 아이들이 si[t]ing 이라고 발음하는 것은 부모의 언어 입력에 나타난 각 변이음(variant)의 상대적인 빈도를 그대로 모방하여 발화한 결과로 볼 수도 있기 때문이다.

아이-어머니 여섯 쌍의 자연발화 말뭉치를 바탕으로, 탄설음화 환경에 있는 어중의 /t/와 /d/가 eating(기본형: eat), hiding(기본형: hide)과 같은 파생어에서 실현되는 양상을 단일어와 비교하여 살펴보았다. 그 결과, 아이들이 파생어의 /t/와 /d/를 발화할 때 기본형에서 사용된 변이음과 똑같은 발음을 사용한다는 것을 알 수 있었고, 이러한 경향은 특히 학습 초기에 나타났다. 구체적으로, 단일어에 포함된 어중의 /t/가 [t]로 실현되는 경우는 거의 없었으나 파생어의 /t/가 [t]로 실현되는 경우는 종종 나타났다. 중요한 것은. 파생어에서 [t]를 발화하는 비율이 어머니들의 언어에 비해 아이들의 언어에서 유의미하게 더 높았다. 이것은 아이들이 어머니의 발화에

나타난 각 변이음의 상대적인 빈도를 자신의 언어에 그대로 적용한 것은 아니라는 점을 시사한다. /d/에 대해서도 비슷한 결과가 나타났다. [d]의 비율은 단일어보다 파생어에서 더 높게 나타났는데, 이것은 아이들이 파생어에서 탄설음을 발화할 때 교체를 회피하려는 경향을 보인다는 것을 의미한다. 어머니들의 발화에서는 단일어와 파생어의 /d/가 [d]로 실현되는 비율이 유의미하게 다르지 않았으므로, 아이들의 언어에서 발견되는 각 변이음의 비율은 어머니의 언어에서 그대로 베껴 온 것이라 할 수 없다. 따라서, 아이들이 기본형과 파생형의 대응되는 분절음을 같은 발음으로 발화하는 이유는 학습 초기에 출력형 대 출력형 충실성 제약이 상위에 위치해 있기 때문이라는 결론을 내릴 수 있다.

본 논문에서는 교체를 회피하는 아이들의 발화 패턴을 설명하기 위해 ID-OO(son)과 ID-OO(voi)를 포함한 최적성 이론 제약을 설정하였다. 이 제약들을 바탕으로 Goldwater & Johnson(2003)에서 제시된 최대 엔트로피 문법(maximum entropy grammar)을 사용하여 아이들의 학습 경로를 모델링하였다.

주요어: 학습 편향, 음운 교체, 미국 영어 탄설음화

학번: 2015-20056