BUCLD 48 Proceedings
To be published in 2024 by Cascadilla Press
Rights forms signed by all authors

A Learning-Based Account of Non-Productivity in Dutch Voicing Alternations

Caleb Belth

A major topic of study in morphophonological acquisition is the development of productive knowledge of alternations. The Dutch voicing alternation, wherein some stem-final obstruents alternate across noun paradigms, has received particular focus. Dutch-learning children appear to acquire the alternation over a protracted period, with little evidence of the alternation being productive by the time children reach 3.5yrs old—an age by which children have usually learned many of the generalizations of their language's morphology. In this paper, I propose a learning-based explanation for this developmental finding. I argue that the distributional properties of the Dutch voicing alternation relegate it to a rather marginal corner of a child's developing lexicon, such that the child may treat alternating forms as exceptions to productive morphological generalizations that are not sensitive the alternation. To make this proposal concrete, I provide a learning algorithm for Dutch noun morphophonology, which demonstrates precisely how these distributional properties can lead to the observed developmental facts. Taken together with recent applications of the learning model to Turkish, the learning model also provides an account of how differences in the distributional contexts of different alternations can account for different developmental trajectories.

1. Introduction

Ever since Berko (1958) introduced the *Wug test*, language acquisition researchers have had an effective tool for assessing the development of productive knowledge of morphophonological generalizations. In some cases, experimental and acquisition evidence suggests that young learners do have productive, rule-like knowledge. For instance, children overextend morphophonological generalizations (e.g., MacWhinney 1978) and apply them to nonce words (e.g., Berko 1958). But in other cases, experimental and acquisition evidence suggests that learners may lack productive knowledge of alternations.

One alternation that has received a wealth of attention is the Dutch voicing alternation. Stem-final obstruents, which are syllabified as a coda in the singular, occur as an onset in the plural, if the plural is formed with the vowel-initial suffix [-ən]. Because Dutch exhibits a phonotactic restriction against voiced obstruents in final position, the resyllabification of stem-final obstruents leads to alternations in some paradigms, like (1a), while other paradigms have voiceless final obstruents throughout, as in (1b) (examples from Zamuner et al. 2006, 2012).

(1) a. [bɛt] 'bed' [bɛdən] 'bed'-PL
b. [pɛt] 'cap' [pɛtən] 'cap'-PL

The Dutch alternation is often accounted for by positing a voiced obstruent in the underlying form of alternating nouns, which is devoiced in final position. For instance, in this account, the underlying form of the noun root in (1a) is taken to be /bɛd/, which is transformed into [bɛt] by a productive devoicing rule. However, acquisition and experimental studies show limited or no evidence of Dutch-learning children applying the voicing alternation productively at age 2;6-3;6 and probably much longer (Zamuner et al. 2006, Kerkhoff 2007, Zamuner et al. 2012). In particular, children have difficulty recognizing or producing the singular form of plural nonces that have stem-final voiced obstruents, such as the singular [slɑt] for the plural nonce [slɑdən].[1]

[1] See § 2 for a detailed discussion.

Final devoicing and its corresponding voicing alternations are common cross-linguistically, yet the developmental trajectory and diachronic fate of the resulting voicing alternations diverges substantially, even among phylogenetically related languages. For example, Yiddish, also a West Germanic language, lost the voicing alternation in nouns to paradigm leveling (King 1976, 1980, Albright 2004). Five-year-old children learning German's analogous voicing alternation sometimes extend it to nonce words (van de Vijver & Baer-Henney 2014)—a possible indicator of productivity. And at 3yrs old, German-learning children appear to have more robust knowledge of the alternating nouns in their lexicon than that of their Dutch-learning peers (Buckler & Fikkert 2016).

One proposal, made clearly by Buckler & Fikkert (2016), is that distributional differences between languages are what lead to the observed differences between developmental trajectories. For instance, Buckler & Fikkert (2016) note that German has a voicing contrast over a larger range of obstruents than does Dutch. Dutch also has two suffixes that form plurals ([-ən] and [-s]), only the first of which creates the environment for voicing alternations, whereas German has five plural suffixes, three of which ([-ə], [-ən], [-ər]) create possible environments for the alternation. Thus, German has a larger number morphophonological environments that have the potential to alternate. Indeed, in a corpus study of child-directed speech, Buckler & Fikkert (2016) found that German exhibits a higher total and relative number of alternating paradigms than Dutch, measured both in type and token frequency.

While developmental differences may be attributable to distributional differences, it is desirable to have a theory about the processes of acquisition to explain how the distributional differences lead to the different developmental trajectories. Recent work has proposed that learners initially construct surface-faithful underlying representations (Tesar 2013), and set up more abstract underlying representations when and only when surface alternation becomes pervasive enough that morphological generalization based on concrete representations becomes untenable (Richter 2018, 2021, Belth 2023a,b). In this paper, I extend this proposal to provide a possible learning-based explanation for how the distributional properties of voicing alternations in Dutch nouns could lead to the apparent non-productivity of the alternation for young learners. The learning model predicts lack of productive generalization of the alternation for learners up to at least several hundred nouns, suggesting that young Dutch learners may lack productivity for this alternation because it forms only a small corner of their developing lexicon and does not prevent morphological generalization from surface-faithful representations. This study helps lay the groundwork for a comparative study of the acquisition of Germanic voicing alternations (§ 5). Moreover, taken together with results from Belth (2023a,b) on Turkish, the model correctly predicts the productivity of Turkish vowel harmony at a young age (Altan 2009) yet the lack of productivity of the voicing alternation in Dutch noun paradigms (Zamuner et al. 2006, 2012).

In § 2 I provide the relevant facts about Dutch noun plurals and the voicing alternation. I then describe my proposed learning model in § 3 and apply it to Dutch noun plurals in § 4. I conclude with a discussion in § 5.

2. The Facts of Dutch Noun Plurals and Voicing Alternations

Dutch, like many other languages (e.g., German and Polish), exhibits a phonotactic restriction against syllable-final voiced obstruents. Beyond the distribution of voiced obstruents, this restriction manifests in the form of alternations in some noun paradigms. Many Dutch plural nouns are formed by suffixing [-ən]. When this plural suffix attaches to a stem ending in an obstruent, the obstruent is syllabified as the onset of the syllable containing the plural suffix. In some noun paradigms—such as (2a; data from Zamuner et al. 2012: p. 482)—the stem-final obstruent is voiced in the plural form, but unvoiced in the singular, where it occurs in syllable-final position. In other paradigms, the obstruent is voiceless throughout—e.g., (2b). Moreover, some noun stems end in a sonorant, in which case no voicing alternation can occur (2c).

(2) a. [bɛt] 'bed' [bɛdən] 'bed'-PL
b. [pɛt] 'cap' [pɛtən] 'cap'-PL
c. [blum] 'flower' [blumən] 'flower'-PL

Because not all nouns or even all nouns with stem-final obstruents alternate, researchers have taken interest in whether Dutch learners are aware of which nouns alternate and which do not (i.e., are they sensitive to the alternation for nouns that they know?), and, if so, whether learners generalize productively so as to expect stem-final voiced obstruents in novel plural nouns to be voiceless in the singular.

In an age-controlled experiment, Buckler & Fikkert 2016 found German 3-year-olds show more advanced knowledge of which nouns in their mental lexicon alternate in comparison to 3-year-old Dutch children. The authors argue that this is likely attributable to language-specific factors, including the fact that German uses voicing contrastively over a broader range of obstruents than Dutch does and that Dutch exhibits both progressive and regressive voicing assimilation while German only commonly exhibits progressive voicing assimilation. These cross-linguistic differences may make the voicing alternating in noun paradigms harder to acquire for Dutch learners than German learners. Perhaps more significantly, Buckler & Fikkert performed a corpus study of German and Dutch nouns in CHILDES (MacWhinney 2000), and found that the analyzed German child-directed speech contained a much greater number of alternating noun paradigms than that of Dutch, whether counted in type frequency (72 in German vs. 22 in Dutch) or token frequency (1572 in German vs. 158 in Dutch).

Other studies have investigated whether knowledge of the Dutch voicing alternation is productive for children, and found evidence that it is not. Zamuner et al. (2006) performed an identification-based reverse wug test with 2.5-year-old and 3.5-year-old Dutch-learning children. A reverse wug test presents the plural form of a nonce noun and asks participants to produce or identify the singular, which flips the classical setting originating from Berko (1958), in which the participants are asked to produce the plural form from the singular nonce. Zamuner et al. (2006)'s study presented children with nonce plurals, together with an image of the nonce entity. Example nonce plurals are in (3). Some nonce words have a stem-final voiced [d] (3a), and thus alternate, while others have a stem-final voiceless [t] (3b), and are thus voiceless throughout the paradigm.

(3) a. [slɑdən]
[kɛdən]
b. [klatən]
[jɪtən]

Following each nonce plural, the experimenter read the singular form and asked the child to point to the correct image. There were three image choices: the plural image of the nonce, the singular image of the nonce (the correct choice), and a distractor (a plural image of a different nonce). For example, a child would be shown a plural picture and told that it represents multiple [slɑdən], and then asked to identify the picture of a [slɑt].

Children show a robust cross-linguistic ability to morphologically decompose nonce words inflected with a familiar suffix, even at the much younger ages of 6-15mo (Marquis & Shi 2012, Mintz 2013, Ladányi et al. 2020, Kim & Sundara 2021) and despite suffix alternations (Ladányi et al. 2020) and in the absence of meaning (Marquis & Shi 2012, Kim & Sundara 2021). The children in Zamuner et al. (2006)'s study indeed showed an ability to identify the singular form of nonce plurals, as they performed above chance at the task overall. But their performance was significantly worse for alternating paradigms like (3a) than for non-alternating paradigms like (3b).

The surface-faithful root form of an alternating nonce like [slɑdən] is [slɑd], which conflicts with the singular form [slɑt] produced by the experimenter. If the children possessed productive knowledge of the voicing alternation, then they should be able to apply it to [slɑd] to eliminate the discrepancy between it and [slɑt]. In that case, there should be no difference between their performance on alternating vs. non-alternating nouns. Thus, the fact that children were significantly worse at identifying singulars of alternating nonce plurals than non-alternating ones suggests that they lack such productive knowledge.

Zamuner et al. (2012) performed a symmetrical production-based reverse wug test with Dutch learners of the same age and with the same stimuli: children were asked to produce the singular of nonce plurals rather than identify it. The results corroborate the identification results from Zamuner et al. 2006, though the children's performance was worse overall on the production task than the identification task.

On the one hand, these results are perhaps not surprising. Buckler & Fikkert (2016) found Dutch 3-year-olds to show little knowledge of which nouns in their mental lexicon alternate; lacking such robust knowledge, there would be little information for children to use to construct a productive generalization for the voicing alternation. Yet, Dutch-learning children almost never produce final voiced obstruents

(Kerkhoff 2007), which suggests that they are sensitive to the phonotactic restriction against them. Evidently, at least by 3;6, this phonotactic awareness has not translated into a productive morphophonological process.

Kerkhoff (2007: ch. 6) found that adult speakers of Dutch were able to produce singulars for nonce plurals from both alternating and non-alternating paradigms, and Ernestus & Baayen (2003) demonstrated that adult Dutch speakers are sensitive to statistical properties of their lexicon when guessing the underlying form of nonce words. Thus, it appears that the voicing alternation may eventually emerge as productive. However, such productivity may be aided by orthography: Kerkhoff (2007) observes that the orthographic rendering of (2a) *bed* ~ *bedden* makes the status of the final [t] in [bɛt] as a devoiced /d/ transparent, and children continue to show little evidence of productively generalizing the alternation even at age 7.

Beyond the distribution described in (2), a number of other distributional facts about the alternation are important to highlight. Some nouns are formed by suffixing [-s], not [-ən], to the stem; (4) gives examples.

(4) [voːɣəl] 'bird' [voːɣəls] 'bird'-PL
[vɪŋər] 'finger' [vɪŋərs] 'finger'-PL

The [-s] plural is usually used after syllables that do not have main stress (Booij 1999: p. 82). Because [ə] does not bear stress in Dutch (Booij 1999: p. 5), a very strong indicator of the [-s] plural suffix is a stem ending in a syllable with a [ə] nucleus. The [-s] suffix is less frequent than the [-ən] suffix, does not often attach to a stem with a final obstruent, and does not lead to a stem-final consonant being resyllabified. Thus, the [-s] suffix does not directly contribute to the Dutch voicing alternation, and nouns taking it are usually ignored in corpus and experimental studies. However, a complete story of Dutch nouns must account for these, and the dataset I construct for this study (see § 4) retains them.

Dutch nouns frequently occur in the diminutive (both singular and plural), with especially common occurrence in child-directed speech. In fact, some plural diminutive forms of a noun are more frequent than their non-diminutive counterparts. For example, Kerkhoff (2007: 113) found that *eendjes* 'ducklings' occurs more times (eight) in her corpus of child-directed speech than does *eenden* 'ducks' (three). This is an important fact because nouns with voiced stem-final segments in the (non-diminutive) plural are unvoiced in both the singular and plural diminutives, as in the (non-diminutive) singular. For example, in (5) the stem-final obstruent of *paard* 'horse' surfaces as [t] in all forms except for the non-diminutive plural (5b).

(5) a. [paːrt] 'horse'
b. [paːrdən] 'horses'
c. [paːrtjə] 'horsie'
d. [paːrtjəs] 'horsies'

If the child learns (5c) 'horsie' and (5d) 'horsies' before learning (5b) 'horses'—a plausible scenario given the prevelance of dimnuitves in child-directed speech—then the child receives no evidence of an alternation. This highlights the importance of considering the details of the child's input in understanding what representations and generalizations they may construct.

Thus, not all Dutch noun plurals are formed by adding [-ən] (4); those that do not will not alternate. Not all plurals formed with [-ən] end in an obstruent (2c); those that do not will not alternate. Not all obstruent-final stems with plurals formed via [-ən] alternate (2b). Even for the subset that do, only 1 of 4 possible forms provides evidence for the alternation (5). These distributional facts form the basis for my proposal: given the sparsity a child's input, the inevitability of exceptions, and the relatively marginal noun environments that exhibit the voicing alternation, children may simply treat the few alternating forms as exceptions to morphological generalizations that are not sensitive to them.

In the remaining sections, I will make this proposal concrete, by providing an explicit learning model for Dutch noun morphophonology. I will argue that the learning model provides a possible learning-based explanation for the apparent lack of a productive morphophonological process corresponding to the voicing alternation for Dutch children at 2;6-3;6.

3. A Learning Model for Dutch Noun Morphophonology

The model used in this study was developed in Belth (2023a,b), building on the models of Richter (2018, 2021) and Belth et al. (2021). I will describe the model background in § 3.1, and then how I extended and applied the model to the question of Dutch voicing alternations in § 3.2.

3.1. Model Background

Because of the Zipfian character of language, most words in linguistic corpuses tend to have low *paradigm saturation*, which measures, for each lemma, the fraction of inflectional categories for that lemma that are attested in the corpus (Chan 2008). This has implications for the empirical conditions of language acquisition because it means that children must generalize productively from their sparse input to the unseen forms of lemmas (see Chan 2008; Yang 2016: ch. 2; Belth 2023b: ch. 2 for further discussion). Belth et al. (2021) proposed an abductive model of this generalization process. The model learns productive morphological generalizations from (*lemma, features, inflection*) triples by recursively subdividing the input until productive generalizations can be formed, where productivity is measured via the Tolerance Principle (Yang 2016). The recursive subdivision yields a decision tree; each leaf of the tree corresponds to a morphological rule, which applies over the subset of the lexicon specified by the path to that leaf from the root. We will turn to this model in more detail in § 3.2.2.

Many lines of inquiry and theoretical approaches have proposed the view that learners' underlying representations are initially surface-faithful. Kiparsky (1968) observed that in the absence of alternation, children have no reason for constructing more abstract representations. Optimality-Theoretic approaches usually posit that children's starting point is to construct underlying forms identical to their observed surface realization (Hayes 2004, Tesar 2013). Ringe & Eska (2013) observe that historical sound changes support the view, which they call *invariant transparency*. Experimental evidence also supports it, as I argue in Belth (2023b: sec. 2.3).

Building on this idea, Richter (2018, 2021) proposed that when surface alternation accumulates into enough violations of *mutual exclusivity*—the principle that each form has a distinct meaning—then morphological generalization from surface-faithful forms can be disrupted, driving the learner to collapse some forms into more abstract underlying categories. Richter uses the Tolerance Principle to measure the precise point at which surface alternation breaks morphological generalization from surface-faithful forms. In particular, Richter carried out a detailed study of the acquisition of the English flap [ɾ] as an allophone of an abstract underlying /T/. Flapping can lead to alternations like [it] 'eat' ∼ [iɾɪŋ] 'eating', which Richter argues lead to enough exceptions to the generalization "if surface [ɾ], then underlying /ɾ/" that learners sometime between 3;0 and 5;0 collapse [t] and [ɾ] into a single underlying category /T/, as evidence by an increased rate of producing *[t] in contexts where [ɾ] would be expected.

Belth (2023a,b) extended this approach from allophones to allomorphs, and applied it to Turkish vowel harmony. In this model, surface alternations in morphemes like the plural suffix [-lɑr]/[-ler] block generalizations that would form plurals by adding only one of the surface-faithful suffixes (e.g., "if PL, then [-lɑr]"). This leads the learner to collapse them into a single abstract category /-lAr/, which then triggers the learning of a phonological rule to predict when /-lAr/ surfaces as [-lɑr] vs. [-ler]

3.2. Extending and Applying the Model to Dutch

I will now describe how I extended the model to Dutch. The model operates in three main steps: (1) it first constructs surface-faithful underlying forms; (2) it then learns morphological inflection rules using these underlying forms; (3) if a rule fails to be productive due to surface alternation, the model collapses the alternating morpheme into an abstract underlying category and learns a phonological rule to account for the alternation.

These steps are the same as Belth (2023a), but the details of step 2 differ. In this paper, I use the model from Belth et al. (2021) to learn morphological inflection rules instead of directly counting the number of words in which a given morpheme occurs as something other than its most frequent form, as I did in Belth (2023a). This is so that the model learns inflection rules for the entire Dutch noun system, as will become more clear in § 3.2.2.

3.2.1. Constructing Surface-Faithful Representations (Step 1)

The input to the model, and hence to step 1, is a set of morphologically-analyzed surface forms. A few examples are in (6), which also exemplify the fact that the learning model does not assume that each (or even most) lemmas in the input will be attested in all of their forms. This is critical, given the sparsity of linguistic input.

(6) [paːrt] 'horse'
[blum-ən] 'flower'-PL
[paːrt-jə] 'horse'-DIM
...

The segmented input reflects children's early ability to relate inflected forms of words to their root forms. Experimental studies have found that 11mo-old French-learning (Marquis & Shi 2012), 6-15mo old English-learning (Mintz 2013, Kim & Sundara 2021), and 15mon-old Hungarian-learning children (Ladányi et al. 2020) can relate inflected forms to their root forms even in the presence of agglutination and alternation.

The model begins, as in Belth (2023a), by constructing a surface-faithful underlying form for each morpheme. The model uses the most frequent surface realization of the morpheme as its surface-faithful UR. For example, because, as described in (5), the root *paard* 'horse' ends in [t] in all forms except the non-diminutive plural, the model will construct /paːrt/ as the UR unless the non-diminutive plural happens to be the only form attested in the input.

The model then attempts to learn morphological inflection rules, which we turn to next.

3.2.2. When is Abstraction Needed? (Step 2)

To learn inflection rules, I use ATP (Belth et al. 2021), which learns from (*lemma, features, inflection*) triples. The present model constructs one such triple for each word in the input. The model uses the word form as the *inflection* of the triple, and looks up the surface-faithful underlying form for the word's root, which was computed in step 1, to use as the *lemma*. The *features* are taken from the word's morphological analysis. For example, the input in (6) are converted into the triples in (7).

(7) ([paːrt] { }, [paːrt])
([blum] {PL }, [blumən])
([paːrt] {DIM }, [paːrtjə])

In addition to the morphological features, the model also adds a phonological feature "[ə] nucleus" to every triple where the root-final syllable has a [ə] nucleus. In principle, this feature can be learned (e.g., by Belth 2023b, To appear), but for simplicity I had the model directly annotate it since, as described in § 2, the feature is relevant for learning whether a noun plural is formed by [-ən] or [-s]. Thus, for example, the input [voːɣəl-s] 'bird'-PL yields the training instance ([voːɣəl], {PL, "[ə] nucleus"}, [voːɣəls]).

The resulting triples serve as the input to ATP. Belth et al. (2021) provides a detailed description of ATP, but I will describe the relevant aspects here. I will use (8) as an example input, which for simplicity includes only plurals. When I apply the model to child-directed corpus data (§ 4), the data will also include singulars, diminutives, and diminutive plurals.

(8) a. ([buk], {PL }, [bukən])
([pɛt], {PL }, [pɛtən])
([blum], {PL }, [blumən])
([vut], {PL }, [vutən])

b. ([bɛt], {PL }, [bɛdən])
([oːx], {PL }, [oːɣən])
([ɦant], {PL }, [ɦandən])

c. ([voːɣəl], {PL, "[ə] nucleus"}, [voːɣəls])
([vɪŋər], {PL, "[ə] nucleus"}, [vɪŋərs])
([lɛtər], {PL, "[ə] nucleus"}, [lɛtərs])
([mudər], {PL, "[ə] nucleus"}, [mudərs])

ATP starts by identifying the most frequent morphological marker (in this case suffix), and proposes a morphological inflection rule to explain it. The most frequent suffix in (8) is [-ən], which occurs in 7/11 word forms. ATP thus proposes the rule "* → [-ən]", which predicts that all inflected forms are formed by adding [-ən]. ATP evaluates proposed rules in terms of the Tolerance Principle (TP) (Yang 2016).

The TP hypothesizes that children form productive generalizations when the number of exceptions to a proposed generalization results in a real-time processing cost lower than that without the generalization. The precise threshold provided by the TP, which is derived in Yang (2016: ch. 3) and incorporates the empirical observation of linguistic sparsity that I discussed above, is given in (9). Here N is the number of items the rule being evaluated applies to and e is the number of exceptions to the rule.

(9)

$$e \leq \frac{N}{\ln N}$$

The TP has had much prior success in computational modeling, lexical, and experimental studies (Schuler et al. 2016, Yang 2016, Richter 2018, Koulaguina & Shi 2019, Emond & Shi 2021, Richter 2021, Belth et al. 2021, Payne 2022, Belth 2023a,b).

The initial rule "* → [-ən]" makes 11 predictions ($N = 11$), 7 of which are exceptions ($e = 7$): the stem-final obstruents in (8b) alternate in addition to adding [-ən], and the items in (8c) add [-s] instead.

When a proposed rule fails the TP test (9), which is the case in our toy example since $7 > 11/\ln 11 = 4.6$, ATP chooses a single feature to subdivide the items based on and then repeats the rule proposal and evaluation procedure recursively over the resulting subdivisions. ATP chooses the feature that is either shared by or absent from the largest number of the words that take the most frequent suffix, ignoring features that are shared by all words. In this case, the feature PL is ignored because it is shared by all words (hence splitting on it would not yield a partition). The feature "[ə] nucleus" is absent from all 7 words that take [-ən] (8a)-(8b), so ATP partitions the lexicon into the groups "[ə] nucleus" and "¬[ə] nucleus", and recurses over those groups.

Within the "[ə] nucleus" group, the most frequent suffix is [-s], so ATP proposes the rule "[ə] nucleus → [-s]", which makes 4 predictions with no exceptions and is thus productive.

Similarly, within the "¬[ə] nucleus" group, the most frequent suffix is [-ən], so ATP proposes the rule "¬[ə] nucleus → [-ən]." This rule makes 7 predictions, and the three alternating forms (8b) can be lexicalized as exceptions since $3 \leq 7/ln7 = 3.6$. Suppose that the input also included the item ([paːrt], {PL }, [paːrdən]). Then the rule would instead have $N = 8$ and $e = 4$, which would break the generalization ($4/8 \ln 8 = 3.8$).

This forms the core of the learning proposal: if, at any point in the process of learning morphological inflection rules, the number of alternations become pervasive enough that they cannot be tolerated as exceptions (as measured by the TP), then the learner must collapse the stem-final obstruents of the offending alternating nouns into an abstract category (e.g., [d], [t] → /D/). If this happens, the formation of the new, non-surface-faithful UR triggers the learning of a phonological generalization to predict the UR's surface realization, via Belth (To appear).

3.2.3. Constructing Non-Surface-Faithful URs (Step 3)

If the voicing alternation becomes pervasive enough that generalization from concrete forms is no longer tenable, the model constructs non-surface-faithful URs for the alternating morphemes. The process for doing so is described in Belth (2023a: sec. 2.5). For brevity I will not re-describe the procedure here, because, as we will see, the need to apply it never arises in the present study (§ 4).

4. Applying the Model to Dutch Nouns

4.1. Setup and Data

To construct a dataset for the study, I extracted nouns from the child-directed speech in the Van Kampen (N. J. Van Kampen 1994, J. Van Kampen 2009) corpus of the CHILDES (MacWhinney 2000) database. I used all sessions with children up to age 3;6, following Buckler & Fikkert (2016)'s corpus study, and the age of the older children in Zamuner et al. (2006, 2012)'s experiments. I used the TreeTagger[2] part-of-speech tagger (Schmid 1999, 2013) to extract nouns from the corpus. The tagger also specifies whether the noun is singular or plural. I tagged diminutive nouns post-hoc based on the *-je* suffix and its allomorphs, which is straight forward to do because the TreeTagger provides lemmas for tagged words. To get IPA transcriptions, I took the intersection of the resulting set of nouns with the Dutch subset of *wikipron* (Lee et al. 2020). Finally, I computed the frequency of each word in the corpus, and dropped words with only a single occurrence. This resulted in a set of 887 nouns—606 singular, 107 plural, 124 singular diminutive, and 50 plural diminutive.

The size of the dataset seems appropriate. Zamuner et al. (2006, 2012) and Buckler & Fikkert (2016)'s experimental and corpus studies involved Dutch-learning children age 2;6-3;6. At around this age, children's vocabularies are consistently measured as falling within a few hundred to a thousand words (Anglin 1993, Fenson et al. 1994, Hart & Risley 1995, Bornstein et al. 2004, Szagun et al. 2006, Kodner 2022). Moreover, nouns tend to be disproportionately represented in children's early vocabularies (Gentner 1982, Bornstein et al. 2004). Thus, a few hundred nouns is likely a reasonable approximation of a Dutch child's vocabulary at 2;6-3;6.

I first ran the model on the entire 887 noun dataset. In § 4.2, I will describe the results and how they provide a possible learning-based explanation for the behavior of 2.5-3.5yr old Dutch-learning children. I then ran the model on multiple, varying-size subsamples of the corpus in order to confirm the robustness of the results (§ 4.3). I also used those simulations to confirm that the proposed model, in addition to providing a possible explanation for 2.5-3.5yr old Dutch-learning children's behavior, also achieves reasonable accuracy predicting novel noun forms, since any plausible morphological learning account must generalize.

4.2. Learning Results

We turn first to the surface-faithful representations constructed by the model in step 1. The nouns *boek* 'book' and *paard* 'horse' both appeared in all combinations in the learning data (SG, PL, DIM, and DIM;PL). The surface-faithful form constructed for *boek* is [buk], and that for *paard* is [paːrt] because they both occur in a majority of noun forms with a voiceless final obstruent (4/4 and 3/4 respectively). Similarly, the noun *eend* 'duck' appeared in the SG, PL, and DIM;PL, but not the non-diminutive plural—consistent with Kerkhoff (2007)'s corpus study—and is thus realized as [eːnt] in all (3/3) noun forms. This demonstrates the relevance of diminutives to the evidence that learners have (or lack) of the alternation. As we will see though, the assumption that learners initially use the most frequent (and hence usually voiceless) form for alternating noun stems is not critical: the productivity conclusions hold up even if the learners use the voiced form (e.g., Albright 2004, in order to account for the direction of paradigm leveling that occurred in Yiddish, suggests that the plural, which is voiced, is privileged in Yiddish noun paradigms).

The output of step 2—i.e. the morphological inflection rules learned by ATP—is shown as a decision tree in Fig. 1. The left-most leaf shows that (non-diminutive) singulars add no suffix (-∅ denotes the null suffix) to the lemma. The [-jə] and [-jəs] leaves show that singular and plural diminutives are formed by suffixing [-jə] and [-jəs] to the lemma, respectively. The remaining two leaves show how plurals are formed: [-s] if the lemma ends in a syllable with a [ə] nucleus, and [-ən] otherwise.

In (10), I have provided the Tolerance Principle counts for each of ATP's rules. The rules are shown as an ordered list from most specific (deepest in tree) to least specific.

[2] https://www.cis.uni-muenchen.de/ schmid/tools/TreeTagger/

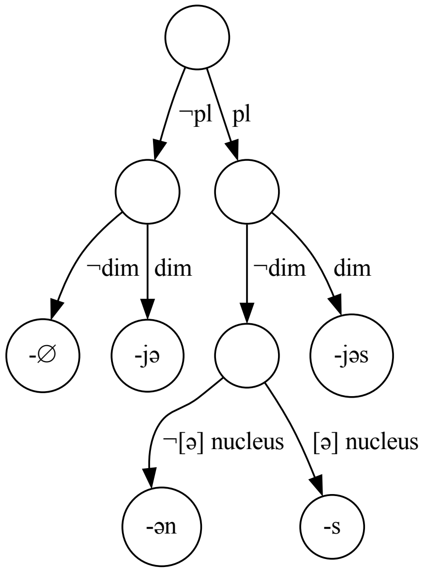

Figure 1: Morphological Inflection Rules

(10)
a. PL ∧ ¬DIM ∧¬[ə] nucleus → [-ən] ($N = 90, e = 10 \leq 90/\ln 90 = 20.0$)
b. PL ∧ ¬DIM → [-s] ($N = 17, e = 2 \leq 16/\ln 16 = 6.0$)
c. PL → [-jəs] ($N = 50, e = 0 \leq 50/\ln 50 = 12.8$)
d. ¬DIM → -∅ ($N = 603, e = 0 \leq 603/\ln 603 = 94.2$)
e. Elsewhere → [-jə] ($N = 124, e = 6 \leq 124/\ln 124 = 25.7$)

Since alternations only occur for paradigms where a stem-final obstruent is resyllabified in the onset of the [-ən] PL suffix, the alternating nouns fall in the third-from-left branch of the tree, whose counts are shown in (10a). Since $10 \leq 90/\ln 90$, the rule is productive despite the alternation, and the alternating nouns can be lexicalized. Moreover, of the 10 exceptions, only 8 are due to the voicing alternation.[3] The parent node of the plural suffixes (10a)-(10b) had $N = 90 + 17 = 107$ nouns, of which 16 take the [-s] suffix and only 8 alternate. Thus, subdividing to isolate the 16 nouns that take [-s] instead of subdividing to isolate the 8 alternating nouns that take [-ən] is expected. Note that if the underlying forms of alternating noun stems contained voiced obstruents instead of voiceless obstruents, the exceptions would instead have gone to the left-most branch (10d) (since devoicing would be necessary to produce the correct singular form), but $8 \leq 94.2$, so the conclusion would hold even more strongly.

It appears, then, that Dutch-learning children may lack productive knowledge of voicing alternations because such knowledge is not critical to their developing linguistic system. The model demonstrates that highly productive morphological generalizations are possible without knowledge of the voicing alternation, and perhaps child learners are content with these imperfect but fully sufficient generalizations.

To make this point more clear, consider the model's behavior on the nonce nouns from the Zamuner et al. (2006, 2012) experiments discussed in § 2. For simplicity, we will consider the identification version of the reverse wug test, in which children were presented with a nonce plural followed by its singular form and were then asked to identify the picture corresponding to the singular. Since the model did not need to learn a voicing alternation, when presented with the plural nonces, the predicted singular form is simply the surface-faithful form of the root in the plural, derived by removing the [-ən] suffix. For example,

[3] The others are due to one stem that takes [-s] and one stem that has an irregular vowel change in the plural.

Table 1: Both humans and the model perform much better at identifying the singular of a nonce plural when the noun does not alternate than when it does. A trigram language model, which captures the phonotactic restriction on voiced final obstruents, does not reflect this asymmetry.

	Non-Alternating	Alternating
Humans	0.61	0.48
My Model	1.00	0.33
Trigram	1.00	1.00

the model predicts [slɑd], [kɛd], [klat], and [jɪt] as the singular for the alternating nonce words in (3), repeated in (11).

(11) a. [slɑdən]
[kɛdən]

b. [klatən]
[jɪtən]

For the non-alternating forms (11b), the model predictions match the singular form spoken by the experimenter. In contrast, the model predictions for alternating forms (11a) do not match the singular form spoken by the experimenter and violate the phonotactic restriction. I model the latter case by supposing that the image is chosen at random (from three choices), in which case the learner will guess the correct form about 1/3 of the time.

The resulting predictions are shown in Tab. 1, where my model predicts the expected form for non-alternating nouns but performs at chance for alternating nouns. For reference, I report the average performance of the children in Zamuner et al. (2006)'s study.[4] Clearly the children performed much worse overall than the model, perhaps indicating the overall difficulty of the task for children. However, the trend is correct. Contrast this with a trigram language model fit to the surface forms of the 887 Dutch nouns. Because voiced obstruents never occur in syllable final position, the [t] singular form spoken by the experimenter always has higher probability than the same noun with a final voiced [d]. This highlights how knowledge of the phonotactic restriction against voiced obstruents in final position does not automatically translate into a productive voicing alternation generalization.

4.3. Quantitative Evaluation

The results above suggest that the voicing alternation is not pervasive enough to require abstract underlying representations on the entire dataset of 887 nouns. To confirm that these observations hold robustly through smaller vocabulary sizes, I simulated earlier stages of development by running the model on the Dutch dataset, simulating incremental learning. Each simulation sampled words with replacement and weighted by frequency, and added them to the lexicon when sampled. In 100-word increments up to a vocabulary size of 700 nouns, I confirmed that productive morphological inflection rules were still extracted via ATP without triggering the collapsing of alternating obstruents into non-surface-faithful URs. In order to confirm that the resulting morphological inflection rules generalize, I also evaluated the accuracy of the morphological inflection rules over 187 test words not seen during training. I performed 30 simulations using different random seeds.

As a baseline, I compared the generalization accuracy to that of a transformer-based (Vaswani et al. 2017) sequence-to-sequence (seq2seq) model. At each training size, I created a random 80/20% train/dev split of the training data, and tuned hyperparameters on the 80% part of the split, using the remaining 20% to evaluate the quality of each hyperparameter combination. I searched the hyperparameters in (12) with the Tree-structured Parzen Estimator (TPE) optimization algorithm via the Optuna Python package

[4] Zamuner et al. (2006)'s results were broken down by age, but no significant effect was found for age, so I aggregated the results into a single number for each type of noun by taking the weighted average of the two age groups' results.

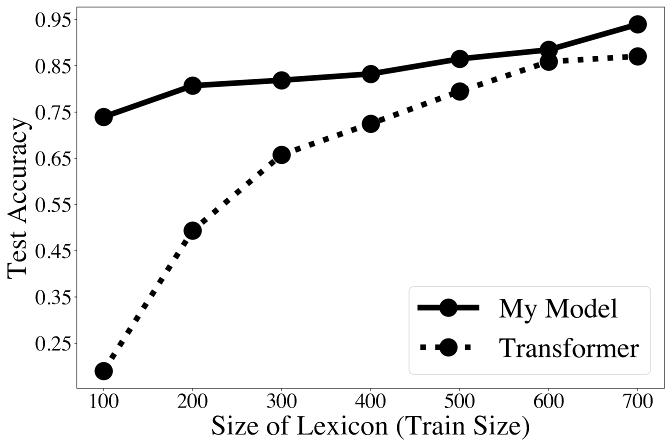

Figure 2: Accuracy generalizing to held-out test words.

(Akiba et al. 2019), and chose the combination with the highest accuracy on the 20% dev set. I then trained a new model with the best hyperparameters on the entire training set (i.e. re-merging the 80/20 split).

(12) learning rate $\in [0.0001, 0.01]$
number of epochs $\in \{10, 11, 12, ..., 29, 30\}$
embedding dimension $\in \{16, 32, 64, 128, 256, 512\}$
hidden dimension $\in \{16, 32, 64, 128, 256, 512\}$
number of attention heads $\in \{1, 2, 4, 8\}$
number of encoder layers $\in \{1, 2, 3, 4\}$
number of decoder layers $\in \{1, 2, 3, 4\}$

Across the 30 simulations, the voicing alternation was never pervasive enough to prevent the model from learning productive morphological inflection rules from surface-faithful representations. Thus, the conclusions from § 4.2 hold up across smaller vocabulary sizes. The accuracy generalizing to held-out test words, shown in Fig. 2, approaches 0.95 by the time the vocabulary contains 700 nouns and surpasses that of the transformer seq2seq model at all training sizes. Despite not learning a productive voicing alternation, the high overall accuracy demonstrates that accurate generalizations are nevertheless possible.

5. Conclusion

While a sensitivity to phonotactics seems to begin developing within the first year of life (Jusczyk et al. 1994, Mattys et al. 1999, Sundara et al. 2022), the acquisition of alternations appears to develop later and the trajectory may vary substantially depending on the alternation (MacWhinney 1978, Tessier 2016). For instance, at least some alternations arising from vowel harmony appear to show productivity before age 3, as has been studied in Hungarian and Turkish (MacWhinney 1978, Altan 2009). This contrasts with the Dutch voicing alternation, which appears to lack productivity up to 3;6 (Zamuner et al. 2006, 2012), probably much longer (Kerkhoff 2007), and perhaps until adulthood (Kerkhoff 2007, Ernestus & Baayen 2003).

The relatively late development of the Dutch voicing alternation could be somewhat surprising, considering the strength of Dutch's phonotactic restriction against final voiced obstruents (Kerkhoff 2007, Tessier 2016). However, distributional properties, such as the frequency of an alternation's attestation, are widely believed to influence the developmental trajectory (Buckler & Fikkert 2016, Tessier 2016). Proposals for the mechanisms that may link the distributional properties to the developmental facts are thus important.

In Belth (2023a,b), I built on ideas and results from Richter (2018, 2021) to propose an algorithmic learning model to account for when surface alternation necessitates non-surface-faithful representations, and hence the construction of a productive (morpho)phonological generalization. The model, when ap-

Table 2: The logic of the model's results across the present study and Belth (2023a). Good generalization accuracy is achievable both when abstraction is needed and when it is not.

Language	Surface-Faithful Generalization?	Abstract URs	High Accuracy
Turkish (Belth 2023a)	✗	✓	✓
Dutch (Present Study)	✓	✗	✓

plied to Turkish (Belth 2023a), demonstrated how vowel harmony, which systemically pervades the morphophonological system, triggers enough surface alternation that effective morphological generalization is only possible if a productive vowel harmony process is learned. These results were consistent with Altan (2009)'s experimental findings regarding the productivity of Turkish vowel harmony in young learners.

The present paper has extended this model to the case of Dutch noun voicing alternations. The model demonstrates mechanistically how a learner might construct morphological generalizations and how that process leads to the few occurrences of voicing alternations being subsumed as exceptions to a fully-productive morphological inflection rule, which simply adds [-ən] to a surface-faithful stem form.

In both Turkish and Dutch—and hence both when it predicts abstraction is necessary and when it is not—the model is indeed able to generalize accurately to novel words. The situation is summarized in Tab. 2. When generalization from surface-faithful URs is not possible (Turkish), the proposed model constructs abstract URs, and high generalization accuracy is achieved by learning to map these to their surface realizations. On the other hand, when generalization from surface-faithful URs is possible (Dutch), the model need not construct abstract URs, but high generalization accuracy is still achieved, directly from the concrete representations.

These results are achieved on descriptively very different alternations. However, as discussed in § 1, even descriptively similar voicing alternations in phylogenetically-related languages appear to have different developmental trajectories and outcomes. Paradigm-leveling led to the loss of voicing alternations in Yiddish (King 1976, 1980, Albright 2004), and German learners appear to develop knowledge of voicing alternations earlier than Dutch learners (van de Vijver & Baer-Henney 2014, Buckler & Fikkert 2016).

My goal is to carry out a comparative study of the acquisition of voicing alternations in West Germanic languages. This goal takes inspiration from Pye & Pfeiler (2014) and Pye (2019)'s *Comparative Method of Language Acquisition*, in which the *contexts of use* of particular structures play a central role and the phylogenetic relationship between languages is taken into account when drawing inferences about the cross-linguistic generality of mechanisms involved in the acquisition of particular languages.

Another direction for future work is to investigate whether my model can shed light on whether orthography is necessary to learning the Dutch alternation, or whether surface alternation becomes pervasive enough to drive its productivity prior to school age. Ernestus & Baayen (2003) demonstrated that adult Dutch speakers appear to use statistical properties of their lexicon to guess the underlying form of nonce words. The study involved verbs rather than nouns, but the results suggest that Dutch learners may eventually extend voicing alternations productively, and that the productivity appears sensitive to relatively subtle distributional properties of the learner's lexicon. It is an open question whether productivity eventually emerges from purely phonological representations, or whether orthography influences the acquisition trajectory. Dutch orthography represents the underlying form of alternating nouns transparently. For example, [bɛt] ∼ [bɛdən] from (1a) is *bed* ∼ *bedden* orthographically. While Barrios & Hayes-Harb (2020) demonstrated that the orthographic equivalence of the stem-final obstruent obscures the auditory alternation for adult learners of a Germanic-like artificial language, to a child learner of Dutch, the transparent representation of the underlying form could lead them to recognize the relationship between alternating forms that they may have missed pre-literacy.

Thus, I plan to investigate the role of orthography by applying my proposed model to incrementally-larger subsets of CELEX (Baayen et al. 1996), as an approximation of the developing vocabulary of Dutch learners. The point at which the model predicts that the alternation should become productive could provide insight into the role of orthography.

References

Akiba, Takuya, Shotaro Sano, Toshihiko Yanase, Takeru Ohta & Masanori Koyama. 2019. Optuna: a next-generation hyperparameter optimization framework. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining* (KDD '19), 2623–2631. Anchorage, AK, USA: Association for Computing Machinery. https://doi.org/10.1145/3292500.3330701.

Albright, Adam. 2004. Sub-optimal paradigms in Yiddish. In *Proceedings of the west coast conference on formal linguistics*, vol. 23, 1–14.

Altan, Asli. 2009. Acquisition of vowel harmony in Turkish. In *35. yıl yazıları*, 9–26. Multilingual Yabancı Dil Yayınları.

Anglin, Jeremy M. 1993. Vocabulary development: A morphological analysis, with commentary by George A. Miller and Pamela C. Wakefield. *Monographs of the Society for Research in Child Development*. i–186.

Baayen, R. H., R. Piepenbrock & L. Gulikers. 1996. The celex lexical database (cd-rom). In.

Barrios, Shannon L & Rachel Hayes-Harb. 2020. Second language learning of phonological alternations with and without orthographic input: evidence from the acquisition of a german-like voicing alternation. *Applied Psycholinguistics* 41(3). 517–545.

Belth, Caleb. 2023a. Towards a learning-based account of underlying forms: a case study in Turkish. In Tim Hunter & Brandon Prickett (eds.), *Proceedings of the society for computation in linguistics 2023*, 332–342. Amherst, MA: Association for Computational Linguistics. https://aclanthology.org/2023.scil-1.29.

Belth, Caleb. 2023b. *Towards an algorithmic account of phonological rules and representations*. University of Michigan dissertation.

Belth, Caleb. To appear. A learning-based account of local phonological processes. *Phonology*.

Belth, Caleb, Sarah Payne, Deniz Beser, Jordan Kodner & Charles Yang. 2021. The greedy and recursive search for morphological productivity. In *Proceedings of the annual meeting of the cognitive science society*, vol. 43. https://escholarship.org/uc/item/1md2p6j5.

Berko, Jean. 1958. The child's learning of english morphology. *Word* 14(2-3). 150–177.

Booij, Geert. 1999. *The phonology of Dutch* (The phonology of the world's languages). Oxford University Press.

Bornstein, Marc H et al. 2004. Cross-linguistic analysis of vocabulary in young children: spanish, dutch, french, hebrew, italian, korean, and american english. *Child development* 75(4). 1115–1139. https://pubmed.ncbi.nlm.nih.gov/15260868/.

Buckler, Helen & Paula Fikkert. 2016. Dutch and german 3-year-olds' representations of voicing alternations. *Language and Speech* 59(2). 236–265.

Chan, Erwin. 2008. *Structures and distributions in morphology learning*. University of Pennsylvania dissertation.

Emond, Emeryse & Rushen Shi. 2021. Infants' rule generalization is governed by the Tolerance Principle. In Danielle Dionne & Lee-Ann Vidal Covas (eds.), *Proceedings of the 45nd annual Boston University Conference on Language Development*, 191–204. Somerville, MA.

Ernestus, Mirjam Theresia Constantia & R Harald Baayen. 2003. Predicting the unpredictable: interpreting neutralized segments in dutch. *Lg* 79(1). 5–38.

Fenson, Larry, Philip S Dale, J Steven Reznick, Elizabeth Bates, Donna J Thal, Stephen J Pethick, Michael Tomasello, Carolyn B Mervis & Joan Stiles. 1994. Variability in early communicative development. *Monographs of the Society for Research in Child Development*. i–185.

Gentner, Dedre. 1982. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In Stan A. Kuczaj (ed.), *Language, thought, and culture*, 301–334. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hart, Betty & Todd R Risley. 1995. *Meaningful differences in the everyday experience of young american children.* Baltimore, MD: Paul H Brookes Publishing.

Hayes, Bruce. 2004. Phonological acquisition in optimality theory: the early stages. *Constraints in phonological acquisition*. 158–203.

Jusczyk, Peter W, Paul A Luce & Jan Charles-Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of memory and Language* 33(5). 630–645.

Kerkhoff, Annemarie Odilia. 2007. *Acquisition of morpho-phonology: the dutch voicing alternation*. LOT dissertation.

Kim, Yun Jung & Megha Sundara. 2021. 6–month–olds are sensitive to english morphology. *Developmental science* 24(4). e13089.

King, Robert D. 1976. *The history of final devoicing in Yiddish*. Indiana University Linguistics Club.

King, Robert D. 1980. The history of final devoicing in Yiddish. *The field of Yiddish* 4. 371–430.

Kiparsky, Paul. 1968. *How abstract is phonology?* Indiana University Linguistics Club.

Kodner, Jordan. 2022. Computational models of morphological learning. In *Oxford research encyclopedia of linguistics*.

Koulaguina, Elena & Rushen Shi. 2019. Rule generalization from inconsistent input in early infancy. *Language Acquisition* 26(4). 416–435.

Ladányi, Enikő, Ágnes M Kovács & Judit Gervain. 2020. How 15-month-old infants process morphologically complex forms in an agglutinative language? *Infancy* 25(2). 190–204.

Lee, Jackson L., Lucas F.E. Ashby, M. Elizabeth Garza, Yeonju Lee-Sikka, Sean Miller, Alan Wong, Arya D. McCarthy & Kyle Gorman. 2020. Massively multilingual pronunciation modeling with WikiPron. English. In *Proceedings of the twelfth language resources and evaluation conference*, 4223–4228. Marseille, France: European Language Resources Association. https://aclanthology.org/2020.lrec-1.521.

MacWhinney, Brian. 1978. *The acquisition of morphophonology* (Monographs of the Society for Research in Child Development). University of Chicago Press.

MacWhinney, Brian. 2000. *The childes project: tools for analyzing talk. transcription format and programs*. Vol. 1. Psychology Press.

Marquis, Alexandra & Rushen Shi. 2012. Initial morphological learning in preverbal infants. *Cognition* 122(1). 61–66. https://pubmed.ncbi.nlm.nih.gov/21937033/.

Mattys, Sven L, Peter W Jusczyk, Paul A Luce & James L Morgan. 1999. Phonotactic and prosodic effects on word segmentation in infants. *Cognitive psychology* 38(4). 465–494.

Mintz, Toben H. 2013. The segmentation of sub-lexical morphemes in english-learning 15-month-olds. *Frontiers in psychology* 4. 24.

Payne, S. 2022. *When collisions are a good thing: the acquisition of morphological marking*. University of Pennsylvania Bachelor's thesis.

Pye, Clifton. 2019. *The comparative method of language acquisition research*. University of Chicago Press.

Pye, Clifton & Barbara Pfeiler. 2014. The comparative method of language acquisition research: a mayan case study. *Journal of child language* 41(2). 382–415.

Richter, Caitlin. 2018. Learning allophones: what input is necessary. In *Proceedings of the 42nd annual boston university conference on language development.* Somerville, MA: Cascadilla Press.

Richter, Caitlin. 2021. *Alternation-sensitive phoneme learning: implications for children's development and language change*. University of Pennsylvania dissertation.

Ringe, Don & Joseph F Eska. 2013. *Historical linguistics: toward a twenty-first century reintegration*. Cambridge University Press.

Schmid, Helmut. 1999. Improvements in part-of-speech tagging with an application to german. *Natural language processing using very large corpora*. 13–25.

Schmid, Helmut. 2013. Probabilistic part-of-speech tagging using decision trees. In *New methods in language processing*, 154.

Schuler, Kathryn D, Charles Yang & Elissa L Newport. 2016. Testing the tolerance principle: children form productive rules when it is more computationally efficient to do so. In *Cogsci*, vol. 38, 2321–2326.

Sundara, Megha, ZL Zhou, Canaan Breiss, Hironori Katsuda & Jeremy Steffman. 2022. Infants' developing sensitivity to native language phonotactics: a meta-analysis. *Cognition* 221. 104993.

Szagun, Gisela, Claudia Steinbrink, Melanie Franik & Barbara Stumper. 2006. Development of vocabulary and grammar in young german-speaking children assessed with a german language development inventory. *First Language* 26(3). 259–280.

Tesar, Bruce. 2013. *Output-driven phonology: theory and learning*. Cambridge: Cambridge University Press.

Tessier, Anne-Michelle. 2016. Morpho-phonological Acquisition. In *The Oxford Handbook of Developmental Linguistics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199601264.013.7.

Van Kampen, Jacqueline. 2009. The non-biological evolution of grammar: wh-question formation in germanic. *Biolinguistics* 3(2-3). 154–185.

Van Kampen, Nelly Jacqueline. 1994. The learnability of the left branch condition. *Linguistics in the Netherlands* 11. 83–94.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser & Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, 5998–6008.

van de Vijver, Ruben & Dinah Baer-Henney. 2014. Developing biases. *Frontiers in psychology* 5. 634.

Yang, Charles. 2016. *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262035323.001.0001.

Zamuner, Tania S, Annemarie Kerkhoff & Paula Fikkert. 2006. Acquisition of voicing neutralization and alternations in Dutch. In *Proceedings of the 30th annual Boston University Conference on Language Development*, 701–712.

Zamuner, Tania S, Annemarie Kerkhoff & Paula Fikkert. 2012. Phonotactics and morphophonology in early child language: evidence from Dutch. *Applied psycholinguistics* 33(3). 481–499.