Phonetic or phonological variation? Learning surface forms for nasalized vowels in a bidirectional OT environment

MA Thesis - Isabel Keijer

University:	Universiteit van Amsterdam
Programme:	Research MA Linguistics
Student number:	10248153
Supervisor:	dr. Silke Hamann
Date:	May 2014

1 Introduction

We investigate the consequences of the adoption of a Gradual Learning Algorithm (GLA) multidirectional Optimality Theoretic (OT) learning model on the assignment of surface forms in allophonic distributions. What factors make learners decide to adopt a branching or non-branching phonological surface structure in this model? In other words, how do learners decide when variation is phonetic or phonological in nature? In phonological theory, allophony is often illustrated by showing a single phoneme that branches into different allophones. The phoneme occurs in different contexts, and can be realised as different allophones depending on the context. This allophonic variation is said to take place in the realm of phonology. However, some coarticulatory/assimilation effects can also be classified as being strictly phonetic. How do learners make the distinction between phonetic and phonological variation? Do learners make the same choices in this distinction when given similar language data as input? Does the proportional distribution of the contrastive elements in their input play a role? We attempt to find answers to these questions by running experimental simulations in the framework of Optimality Theory, assuming a multidirectional GLA learning model.

In the following sections we introduce the concepts and theories central to this research: the bidirectional approach (Section 1.1), Optimality Theory (Section 1.2), the Gradual Learning Algorithm (Section 1.3), multidirectional error detection (Section 1.4), the difference between phonetic and phonological variation (Section 1.5), and vowel nasalisation (Section 1.6). An overview of the research questions is given in Section 1.7. The methods used in this research are described in Section 2. Results are shown in Section 3 and discussed in more detail in Section 4. Section 5 contains our conclusions and ideas for further research.

1.1 Bidirectional approach

In this thesis we adopt the bidirectional approach to phonology and phonetics following Boersma (2007), Apoussidou (2007), Boersma & Hamann (2008), Hamann (2009) and Boersma (2011). In most language models only the production direction is described, but as the name suggests, in a bidirectional model both the production direction and the perception direction are included. Because of the bidirectionality of the model, it allows the role of the listener in phonological processes to be taken into account as well as the role of the speaker, as opposed to traditional speaker-oriented models (see Hamann 2009). The same constraint set (grammar) is used for both production and comprehension. Another feature of the model in Figure 1 is that the phonetic level is included. In traditional phonological theories this phonetic level is usually left out
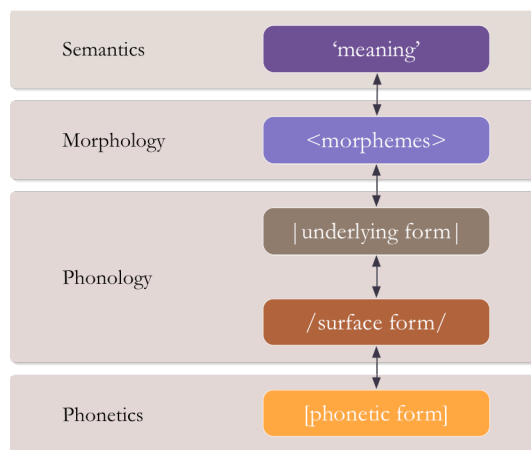
Figure 1: Levels of representation in the bidirectional model

because phonological phenomena are considered separate from phonetics. In including the phonetic level, the bidirectional model takes a more holistic approach to linguistics, leaving room for interaction between phonetic, phonological, morphological and semantic language phenomena. Constraints on different levels of the model interact and are evaluated in parallel. Figure 1 shows the levels of representation that are needed to perceive and produce language in this model. The model in Figure 1 can be read top-down to show the production

process: here, the speaker starts with an intended meaning and then computes the appropriate morphemes, underlying forms, surface forms and phonetic forms. The perception process takes place in the opposite direction (bottom-up): the listener starts with a phonetic form and computes the surface form, underlying form, morphemes and the meaning from that initial phonetic form. As can be seen in Figure 1, the bidirectional model used in this thesis includes two phonological forms: an underlying and a surface form. In her proposal for a listener-oriented model of sound change, Hamann (2009:8) explains that the underlying form "includes only the information that has to be stored in the lexicon", while the lower-level surface form "contains predictable information like foot structure and stress", and is connected to the phonetic form. We henceforth equate the underlying form to the traditional phoneme.

Our focus in the current paper is on the bottom three representations in Figure 1: the phonetic form, the surface form, and the underlying form. Our aim is to see when variation in a learner's input is registered as a phonetic alternation (variation in phonetic forms) or as a phonological alternation (variation in surface forms) in cases with the same underlying form (the difference between phonetic and phonological variation is further discussed in Section 1.5).

1.2 Optimality Theory

The main framework adopted in this thesis is that of Optimality Theory (Prince & Smolensky 1993/2002; henceforth OT). In OT, language users compute optimal input-output pairs through their internal grammar. This means that in production, the speaker starts with an intended meaning as input and uses his grammar to compute the optimal phonetic form as an output. OT is traditionally only used as a model for the production process, not for perception. However, we can also apply OT to perception using the bidirectional phonology model described in Section 1.1.

Table 1 shows the mechanism that is used to compute optimal candidates in OT, a tableau. It consists of the input (top left) and a number of possible output candidates below it. The specific ranking of constraints in the grammar (top right) determines the candidate that is chosen as output. The ranking of the constraints in Table 1 is indicated by the order in which they are shown: constraint C_1 is ranked highest, and constraint C_4 is ranked lowest. In the OT tableau, an asterisk (*) is used to mark if the candidate violates the constraint in that column, so we can see that output candidate 1 violates constraint C_1. This particular violation is also marked with an exclamation mark (!) because this violation leads to the candidate being ruled out: this candidate has violated a top-ranked constraint, while there are still other possible candidates who do not. Once a candidate has been ruled out, the lower-ranked constraints no longer need to be considered for that candidate (in Table 1 the areas that are irrelevant for the decision of the winning candidate are shaded). This means that the total number of constraint violations for a candidate is not what matters; only the highest-ranking constraint violation has an influence in choosing the optimal candidate. Once constraint C_1 has been evaluated and output candidate 1 has been ruled out, the next constraint in the ranking, C_2 is considered. Output candidate 3 violates this constraint, while there is still another possible candidate that does not violate either C_1 or C_2. That means that candidate 3 is eliminated (see the *! mark in the tableau), and only one possible candidate remains. Output candidate 2 is chosen as the optimal candidate (see the pointing finger mark: ☞).

Table 1: Example OT tableau with four constraints

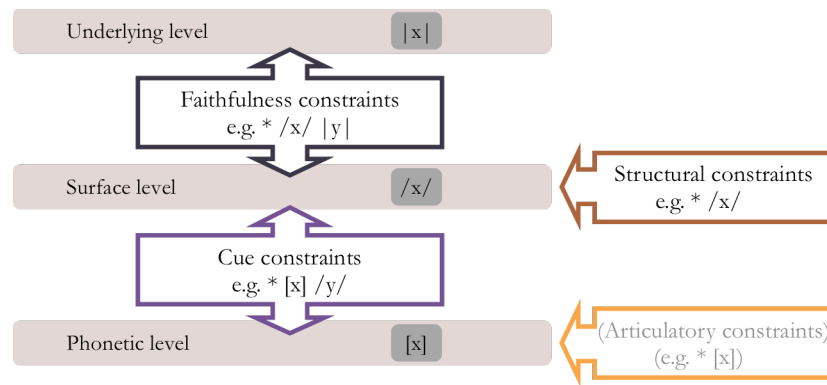| Input	C_1	C_2	C_3	C_4
Output candidate 1	*!		*	
☞ Output candidate 2			*	*
Output candidate 3		*!		*

Figure 2: Constraint types on the levels of representation of phonetics and phonology

Figure 2 shows the types of constraints that play a role in phonological mapping. Faithfulness constraints govern the mapping of forms from the surface level to the underlying level. An example of a faithfulness constraint would be */x/|y|, which means that the surface form /x/ should not be mapped to underlying form |y|. Structural constraints are concerned with surface representations. These structural constraints are used to convey if a certain structure is accepted in the language (or more precisely, if a structure is not accepted; since constraints are always formulated negatively). For example, the structural constraint */x/ means that /x/ is not a valid structure in a grammar. Cue constraints govern the mapping of phonetic forms to surface forms. An example of a cue constraint is *[a]/b/, which means that the phonetic form [a] should not be mapped to surface form /b/. Articulatory constraints deal only with the phonetic forms. Articulatory constraints convey an aversion to forms that require more than average articulatory effort in production. All of these constraints are not to be taken as absolute rules. Instead, the constraint's ranking in relation to the other constraints is what influences the final outcome (see Table 1).

1.3 Gradual Learning Algorithm

In OT, to learn a language is to learn its constraint ranking. Boersma's *Gradual Learning Algorithm* (1997; henceforth GLA) describes how exactly these constraint rankings are learned. In the GLA, a language learner starts with a set of constraints configured in an initial ranking (for instance, all constraints start at the same ranking). The learner is then presented with language data (e.g. by hearing someone speak), and incorporates the evidence from this data into his or her grammar. The constraints in the learner's grammar are re-ranked in response to language input data, but only when "the input data conflict with [the grammar's] current ranking hypothesis"(Boersma & Hayes 2001:45). This means that learning in the GLA is error-driven: learning takes place when an error (i.e. inconsistency with language data) in the learner's current grammar is detected. We

Table 2: Learner's output is different from 'correct' output (adapted from Boersma & Hayes 2001:52)

/underlying form/		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
✓	Candidate 1 ('correct' output)	*!	**	*		*			*
☞	Candidate 2 (learner's output)		*	*	*		*		*

Table 3: Constraints that show a difference in violations between learner's output and 'correct' output are adjusted (adapted from Boersma & Hayes 2001:53)

/underlying form/		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
✓	Candidate 1 ('correct' output)	*→	*→			*→			
☞	Candidate 2 (learner's output)				←*		←*		

Table 4: Constraint ranking values before adjustment

		107	106	105	104	103	102	101	100
	/underlying form/	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
✓	Candidate 1 ('correct' output)	*→	*→			*→			
☞	Candidate 2 (learner's output)				←*		←*		

Table 5: Constraint ranking values after adjustment

		106.8	105.8	105	104.2	102.8	102.2	101	100
	/underlying form/	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
✓	Candidate 1 ('correct' output)	*→	*→			*→			
☞	Candidate 2 (learner's output)				←*		←*		

Table 6: Constraint ranking values after 3x the same adjustment

		106.4	105.4	105	104.6	**102.6**	**102.4**	101	100
	/underlying form/	C_1	C_2	C_3	C_4	**C_6**	**C_5**	C_7	C_8
✓	Candidate 1 ('correct' output)	*→	*→				*→		
☞	Candidate 2 (learner's output)				←*	←*			

discuss the exact way in which we assume these errors to be detected in the learner's grammar in more detail in Section 1.4. In GLA, once an inconsistency between the learner's output (generated by his current grammar) and the 'correct' output has been detected, the constraint violations of the learner's current output and the 'correct' output are compared, and the constraints that show a difference in violations are re-ranked (see Table 2 and Table 3). As Table 3 shows, constraints that are violated exclusively by the learner's chosen output candidate are given a higher constraint ranking. Conversely, constraints that are violated exclusively by the 'correct' output, the output that should have been chosen by the learner, are given a lower constraint ranking. The fact that in GLA constraints are both demoted (ranked lower) and promoted (ranked higher) as a result of the error-detection makes it different from other OT learning strategies such as *Constraint Demotion* (Tesar 1995, *et seq.*).

Another important aspect of the GLA learning process is that the way in which constraints are adjusted is gradual. That is to say, constraints are not immediately promoted or demoted in relation to the other constraints. Instead, each constraint has a numerical ranking value and this ranking value is slightly adjusted each learning cycle. When presented with enough evidence, the categorical rankings will also change. The quantity by which constraints' ranking values are adjusted each time is called the plasticity. As an example of the ranking value adjustment, see Table 4, Table 5 and Table 6. The plasticity in this example is set to 0.2. Table 4 shows some example ranking values with the same situation as in Table 2 and Table 3. Table 5 shows the adjusted ranking values after one learning cycle. We can see that after one learning cycle, the ranking values for constraints C_1, C_2, C_4, C_5 and C_6 have changed slightly, but the categorical ranking is still the same (C_1 is still ranked above C_2, etc.). In Table 6, we can see what happens if the learner is then presented with the same data another two times: the ranking values for C_1, C_2 and C_5 have decreased further while those for C_4 and C_6 have increased further. These continued adjustments have resulted in a different categorical ranking: in Table 6, C_6 is now ranked higher than C_5.

The example above demonstrates the gradual learning process with evaluation noise set to 0 instead of PRAAT's default setting of 2. Evaluation noise is another of the key features of GLA: each time the learner receives an input and wants to evaluate the possible candidates, random noise (normally distributed with the standard deviation being the 'noise' setting, so a default σ of 2) is temporarily added to the ranking values. The resulting disharmony values (i.e. ranking value + noise) are the values that are

actually evaluated to decide the winning candidate. This stochastic component of evaluation noise has the effect that when ranking values are close to each other, the grammar can produce variable outputs and might cause further learning (Boersma & Hayes 2001:46).

The learning cycles described above are repeated until the pre-specified total number of inputs has been fed to the learner. If learning has stabilised by this time, we suppose that the learner has achieved an adult grammar. If learning does not stabilise in time, the learner was unable to form an adult grammar.

1.4 Multidirectional error detection

In this paper we apply a multidirectional error detection method to the learning algorithm of GLA. It has been shown that with such a multidirectional OT learning method, the learner can learn how to compute surface forms and phonetic forms with underlying forms as input, if they are given enough informative pairs of phonetic and underlying forms during learning (Tesar & Smolensky 1998). This method was first applied in a Constraint Demotion learning algorithm by Tesar & Smolensky (1998), called *Robust Interpretative Parsing*. An important feature of this method is that the learner can learn to assign hidden structure to overt (phonetic) forms without having received direct evidence/feedback for that hidden structure. The Robust Interpretative Parsing was later adapted by Boersma (2003) and Appousidou (2007) to incorporate the GLA learning method. In her work on metrical phonology, Apoussidou (2007) expanded on the idea of a feedback loop where the learner evaluates their own hypothetical output in a situation where the learner only receives evidence in the language data for certain forms. Apoussidou applied this process mostly on the semantics-phonology interface, where the phonetic forms were directly discernable from the phonological form.

In this paper we focus on the mapping from phonetic to phonological (surface and underlying) forms, where the same process applies to a situation where the learner perceives a phonetic form, and knows which underlying form is associated with that phonetic form (note that this means we presuppose lexical knowledge of the form). The learner does not know which surface form should be assigned to the phonetic or underlying form. Whereas in the typical OT learning process the learner would receive feedback in the form of a 'correct candidate' in order to learn, our learner that uses multidirectional error detection receives no such evidence for any correct surface form. In comparing both perception and production, this method incorporates the bidirectionality of the approach described in Section 1.1 into the OT learning process. This reflects the reality of language acquisition more accurately than in an entirely supervised approach, since abstract representations are not usually readily available to the L1 learner. In short, multidirectional error detection has the advantage that, apart from the initial language data (an overt phonetic form and a lexical underlying form) the learner receives, there need not be an external factor in the learning process. The learner can evaluate his/her own output and adjust their grammar accordingly.

In Table 7 an example of multidirectional learning is given. In this example, the learner receives a pair of a phonetic form and an underlying form from the

Table 7: Bidirectional error-detection with partial tree input

[x] |y|	104	103	102	101	100
	C_1	C_2	C_3	C_4	C_5
[x] /x/ |x|	*				
[x] /y/ |x|		*		*	
[x] /x/ |y|	*		*		
☞ [y] /x/ |y|				*	←*
✓ ☞ [x] /y/ |y|			*→	*	
[y] /y/ |y|		*	*		*

language data: [x] |y|. This means that in the learner's grammar, the phonetic form [x] should be associated with the underlying form |y|. The learner then generates an output path with the phonetic form as input (perception - marked with an index finger pointing right: ☞), based on his current grammar: [x]/y/|y|. The learner also generates an output path with the underlying form as input (virtual production - marked with an index finger pointing left: ☜): [y]/x/|y|. Lastly, the learner generates an output path that has both the phonetic and underlying form: [x]/y/|y|. The output path that has both the correct phonetic and underlying form is treated as the correct output. The perception output and virtual production output are compared with the combined/correct output, and if they resulted in different output (if an error has been detected), the GLA adjustment procedure discussed in Section 1.3 is applied. In Table 7, the perception output path (☞) is the same as the combined/correct output (✓), so no constraints need to be adjusted for this. However, the virtual production output path (☜) is different from the combined output path, so the GLA adjustment procedure is triggered: all constraints that are violated by the virtual production candidate but not by the combined/correct candidate (i.e. C5) are moved up the ranking hierarchy by one step (i.e. the plasticity value), and all constraints that are violated by the combined/correct candidate but not by the virtual production candidate (i.e. C3) are moved down the ranking hierarchy by one step.

1.5 Phonetic and phonological variation

The goal of this thesis is to investigate how learners react to phonetic variation when assigning surface forms, assuming the framework sketched above in Sections 1.1-1.4: an OT multidirectional GLA framework. We hypothesise that when confronted with variation in the phonetic input, the learner has two ways of categorizing that variation (provided there is no phonemic contrast): the variation can be purely phonetic (see Figure 5), or the variation can be phonological in nature (see Figure 6).

In traditional phonological approaches, there are only two levels of representation: the phonemic (here underlying) level and a level that more or less corresponds to a combination of the phonetic and surface levels that were introduced in Section 1.1. In order to distinguish this level from the ones in the bidirectional model we will refer to it as the 'concrete level' (in contrast to the more 'abstract' phonemic representation). Since traditional phonology approaches exclude phonetics altogether, the focus here is more on phonological forms. Within this traditional approach, distinguishing between phonological and phonetic variation is fairly simple: if variation results in a difference in meaning, the variation is phonological (i.e. there is a difference in phonemes, see Figure 3), and if the variation does not influence the meaning of sounds it is merely phonetic variation (see Figure 4). In this traditional approach we cannot distinguish between allophony and phonetic variation, as both are represented by the split on the concrete
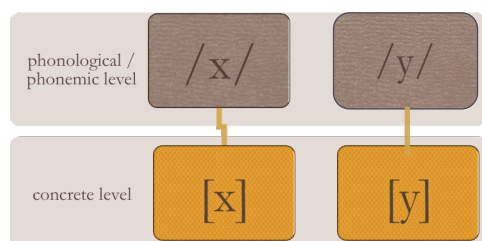
Figure 3: Phonological variation in a traditional two-level phonological approach
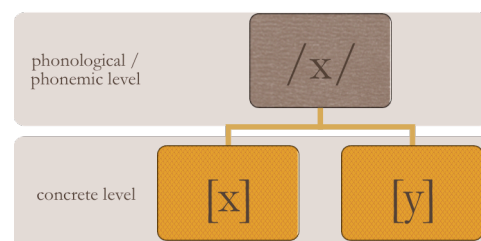
Figure 4: Phonetic variation in a traditional two-level phonological approach

level shown in Figure 4. As has been explained in Section 1.1, in the bidirectional model adopted in this thesis (shown in Figure 1) there are not two, but three levels of representation within the realm of phonology and phonetics: the phonetic level, the (phonological) surface level, and the (phonological) underlying level. In this model, there are more factors at play in distinguishing between phonological and phonetic variation. If there is variation in sounds, this variation can be either:

- Phonological/phonemic (on the underlying level), when the variation results in a difference in meaning;
- Phonological (on the surface level), when there is no difference in meaning, but the variation is categorical and results in allophony, see Figure 6;
- Phonetic, where there is a difference in the auditory input that does not result in a difference in meaning, and variation is gradual/continuous (not categorical), see Figure 5.

Research by Apoussidou (2007) has shown that, when given a sufficient number of informative underlying and phonetic form pairs, learners can also learn to compute surface forms in a multidirectional GLA model. What we don't know is if and how they can learn the distinction between phonetic and phonological surface variation described above. When presented with underlying-phonetic form pairs with variation in the
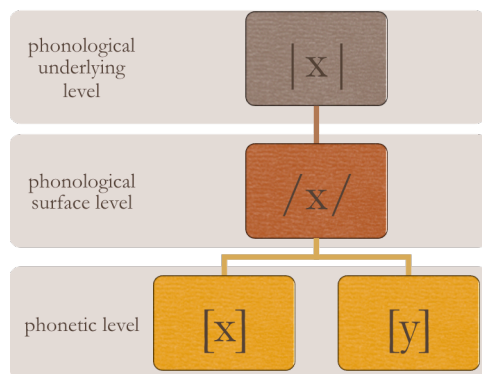
Figure 5: Variation on the phonetic level (non-branching on the phonological level)
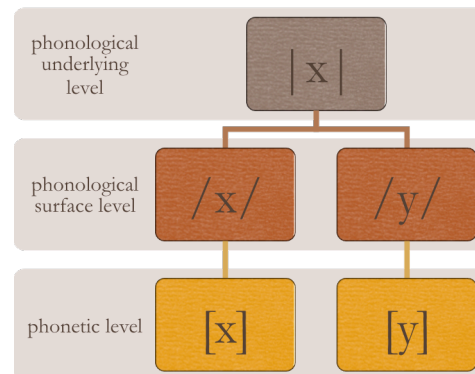
Figure 6: Variation on the surface level (branching on the phonological level)

phonetic forms, will they assign non-branching (Figure 5) or branching (Figure 6) surface forms? What factors influence this choice in assignment strategy?

1.6 Vowel nasalisation

In order to investigate the topic of phonetic and phonological variation introduced in Section 1.5, we have chosen an example of coarticulatory variation: that of vowel nasalisation.

In some languages, vowel nasalisation is phonemically contrastive (e.g. French /pɛ̃/ *pain* 'bread' vs. /pɛ/ *paix* 'peace' [Hajek 2013]). In those languages, there is a difference on the underlying level between nasalised phonemes and non-nasalised phonemes, and this difference in phonemes leads to a difference in meaning between words. In the current paper, we investigate languages where vowel nasalisation occurs, but is *not* phonemically contrastive. Instead, the contrast between nasalised and non-nasalised vowels takes place on the surface and/or phonetic level (see Figure 7 and Figure 8).
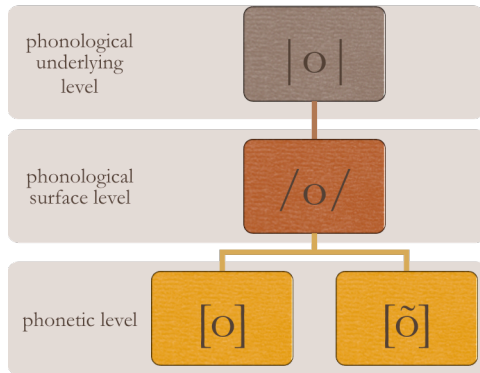
Figure 7: Variation on the phonetic level (non-branching on the phonological level)
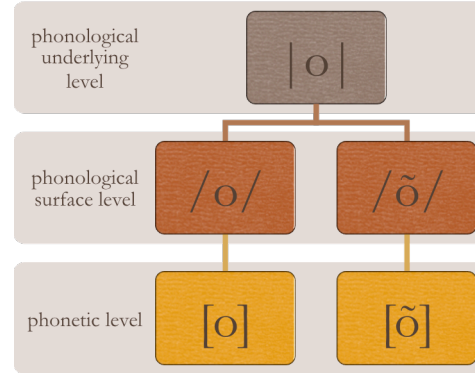
Figure 8: Variation on the surface level (branching on the phonological level)

According to Botma (2004:112), "in languages in which nasalized vowels are derived, these vowels are usually nasalized by a neighbouring nasal consonant," so a form of coarticulation. Kühnert & Nolan (1999:7) explain that the term coarticulation "refers to the fact that a phonological segment is not realized identically in all environments, but often apparently varies to become more like an adjacent or nearby segment". In the case of vowel nasalisation, the segment that is affected is the vowel, which becomes nasalised in the presence of an adjacent nasal segment. Coarticulatory effects can be either anticipatory (the nasal segment follows the affected segment), or progressive (the nasal element precedes the affected segment). We have chosen to only cover the coarticulatory phenomenon of anticipatory nasalisation, since cases of anticipatory vowel nasalisation have received more thorough documentation than their progressive counterparts (for cases of progressive nasalization, see Scheurup 1973).

If we now apply the general issue of distinguishing between phonological and phonetic variation described in Section 1.5 to a concrete language phenomenon, here vowel nasalisation, we arrive at the following problem statement:

Assuming a vowel is (nearly) always nasalised when it precedes a nasal consonant while it is not nasalised in other contexts, how do we know if this is strictly phonetic variation or if this variation is due to two different allophones? What aspects of the variation could play a role in this distinction?

In this paper we investigate only a distinction between two phonetic forms: nasalised or non-nasalised. This is a simplification of the situation in reality, where the degree of nasalisation on the phonetic level is more gradual, and phonetic forms can be partially nasalised. However, the focus of the current research is on the surface form, which is why we have concentrated our efforts on the effects of input distributions, constraint sets and constraint rankings on surface form choice, instead of on phonetic details like partial nasalisation which may also play a role. Our focus is thus more on the quantitative aspects of variation instead of on the qualitative (phonetic) aspects: we hypothesise that in situations where one phonetic form occurs much more often than the other, this variation is more likely to be phonetic, i.e. the same surface form would be used for both phonetic forms. Conversely, our hypothesis is that in situations where both phonetic forms occur in equal (or similar) numbers a learner will perceive the variation as structural and incorporate this structure into the surface level by assigning separate surface representations to the two phonetic forms.

1.7 Research questions

We have formulated the following research questions in our investigation of surface form assignment for nasalised vowels:

Research questions:
1. Is the learners' assignment of surface forms influenced by the presence/absence of symmetrical faithfulness constraints between the surface form and underlying form?
2. Do learners, given the same input distribution in the learning phase, assign the same surface forms to phonetic forms, or is there great variation between learners?
3. Do learners favour one type of surface form assignment (i.e. branching vs. non-branching) over the other?
4. Is the learners' assignment of surface forms influenced by the input distribution of phonetic forms in the learning phase?
5. In what way does the initial ranking of cue constraints influence the learners' assignment of surface forms to phonetic and underlying forms?

2   Methods

In order to answer the research questions formulated in Section 1.7 we have modelled a number of experiments in PRAAT using the GLA (see Section 1.3) in a multidirectional OT (i.e. MultiOT) environment (see Section 1.4). In the experiments 1,000 learners are each given a total of 800,000 partial inputs (no surface forms are given). Based on this language input, they re-rank the constraints in their grammars to form their final (adult) grammar. This learning phase of the experiment is described in further detail in Section 2.1. After learning has finished, we move on to the testing phase. During this phase we ask each learner to generate 10 complete output paths for each possible perception and production input. We ask learners to generate multiple output paths so we can see if the learner is able to generate consistent outputs from its newly learned grammar. In section 2.2 we discuss the testing phase in more detail. In Section 3 we discuss the results of these experiments given different constraint sets (Section 3.1), given different initial constraint rankings (Section 3.2) , and given different input distributions (Section 3.3).

2.1 Learning phase

During the learning phase, each learner is given a total of 800,000 inputs as language data that they can use to form an adult grammar. In the following sections we explain how this language data is structured, and how the learner constructs its adult grammar. For an overview of the learning phase, see Figure 9.
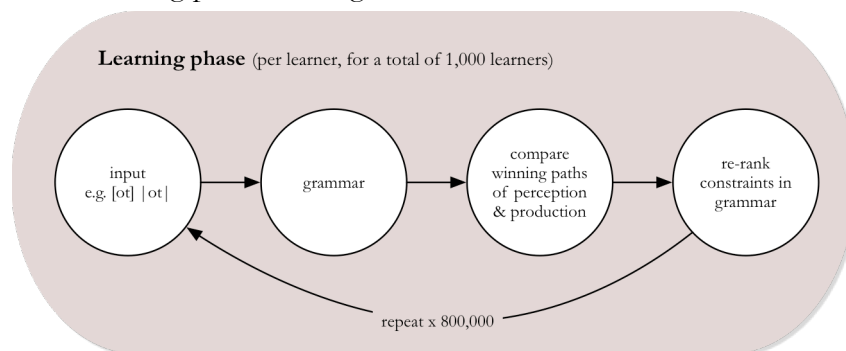
Figure 9: Schematic overview of the learning phase in the experiments

2.1.1 Input

The input received by the learner consists of pairs of phonetic and underlying forms. The phonetic forms consist of the affected vowel (nasalised or non-nasalised) and a nasal ([n]) or non-nasal (here [t]) following consonant. The close-mid back rounded vowel [o] was chosen as symbol for the affected vowel, but in theory any other vowel could be substituted (however, some vowels have been shown to be more susceptible to nasalisation than others, see Young et al. 2001). The same applies to the voiceless alveolar plosive [t], which is used as a symbol for any non-nasal context. We have chosen to provide the learner with some context in the form of the following consonant because of the large role of context in coarticulatory nasalisation.

The inputs that learners receive during the learning phase take the form of partial paths (see Section 1.4): the learner has access to the phonetic form and the underlying form, but not to the surface form. For an overview of all possible partial input paths see Figure 10.
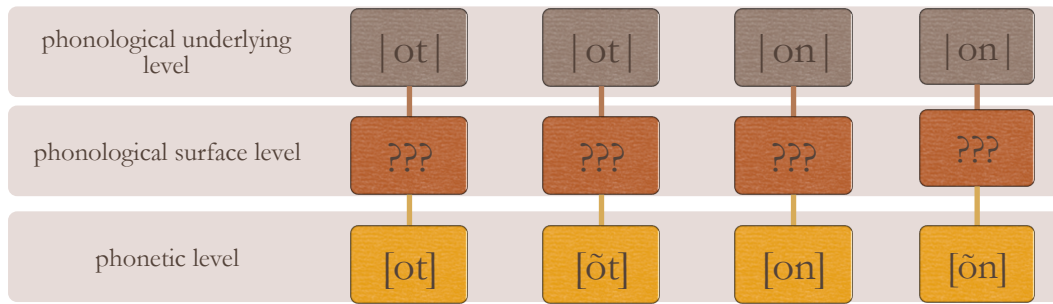
Figure 10: Partial inputs

2.1.1.1 Input distributions

The language data that a learner receives during the learning phase consists of a randomly ordered set of 800,000 inputs. Not all learners receive the same language data. In order to test the influence of input distributions on the learners' choice of surface form (research question 4 in Section 1.7), three different scenarios were tested:

1. learners are exposed to only one phonetic form;
2. learners are exposed to two phonetic forms in equal distribution;
3. learners are exposed to two phonetic forms, of which one is prevalent.

The phonetic forms in the scenarios have been selected for actual occurrence, so as to make the results more easily interpretable. For instance, it is more likely for a nasalised vowel to occur in a nasal context (i.e. [õn]) than for it to occur in a non-nasal context (i.e. [õt]). Consequently, the two phonetic forms in scenarios 2 and 3 are [ot] (non-nasalised vowel with no-nasal context) and [õn] (nasalised vowel with nasal context), as opposed to [õt] and [on]. A bit of noise is added to these scenarios, so some very small evidence of non-standard/prevalent forms (i.e. [õt] and [on]) is present as well, which results in the three input distributions shown in Table 8 and Figure 11.

Table 8: Input distributions

	[ot] |ot|	[õt] |ot|	[on] |on|	[õn] |on|
Scenario 1	97%	1%	1%	1%
Scenario 2	49%	1%	1%	49%
Scenario 3	69%	1%	1%	29%
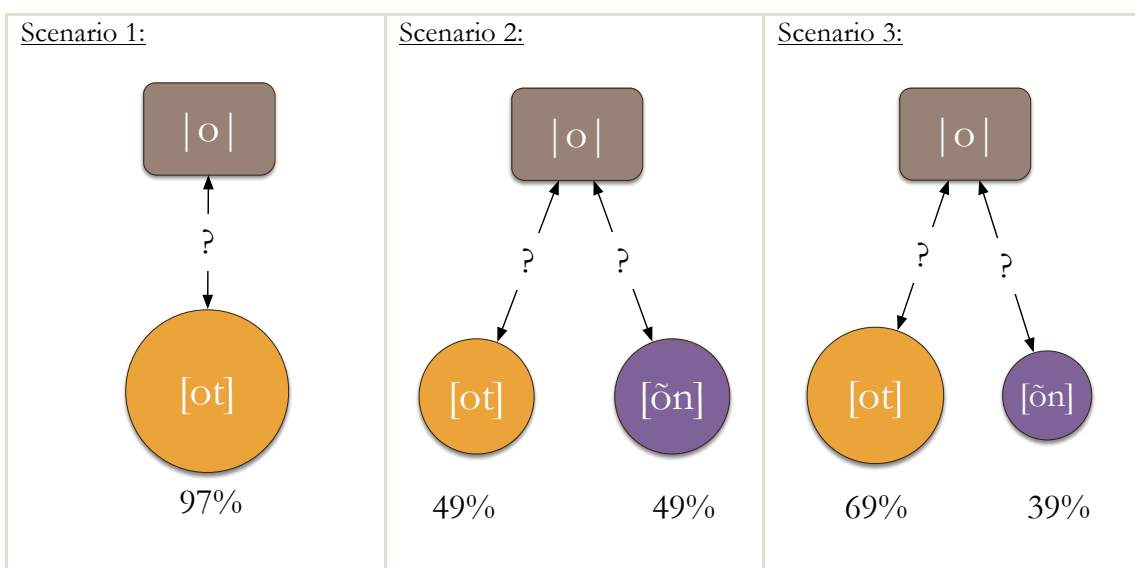
Figure 11: Input distributions

2.1.2 Constraint sets

As was shown in Figure 2, there are four types of constraints that act on the phonetic, surface and underlying levels: articulatory constraints, cue constraints, structural constraints and faithfulness constraints. However, our learners' grammars do not contain any articulatory constraints. This choice was made because of the structure of our learning scenarios: we hypothesise that the learner receives language data through perception (so initial input consists of a phonetic form), and the learner knows which lexical (underlying) form is supposed to be conveyed. The learner forms a hypothesis on the basis of his current grammar as to which candidate path is optimal for the phonetic form he has received as input. The multidirectional feedback loop (or virtual production) then starts, and the learner also generates an optimal output path for the connected underlying form as input, and one optimal output path with both the phonetic and underlying form as input. These paths are then compared and the grammar is altered according to the method explained in Sections 1.3 and 1.4. Articulatory constraints are only relevant in production, since they code for difficulty of articulation. We posit that difficulty of articulation is irrelevant in the learning process described above, since this process is rooted in perception, and the production process involved is purely virtual as part of the feedback loop. Since no actual articulation takes place, difficulty of articulation should not play a role in learnability of a perceived phonetic form.

The total constraints included in the learners' grammars are shown in Table 9 and in Figure 12. The grammar's cue constraints include constraints against all possible combinations of phonetic and surface forms, and structural constraints include constraints against both possible surface forms for the two contexts. As is shown in Table 9, we test two different constraint sets that are identical except for the faithfulness constraints. Usually faithfulness constraints are only formulated as constraints against a coupling of two representations that are not 'faithful' or identical in some aspect, i.e. */x/ |y|. This is the case for the faithfulness constraint in set A. However, this makes the constraint set asymmetric (because there is no opposite constraint) which we predict may have an influence on our results. In order to negate any possible influence of asymmetric constraints we have also designed the symmetric constraint set B (see Table 9). We use the two different constraint sets, A and B, in order to answer research question 1 formulated in Section 1.7, which concerns the influence of symmetric and asymmetric constraint sets on the assignment of surface forms.

Table 9: Constraint sets A and B

Constraint type	Constraint set A (asymmetric)	Constraint set B (symmetric)
Faithfulness constraints	*/õ/ |o|	*/õ/ |o|
		*/o/ |o|
Structural constraints	*/on/	*/on/
	*/õn/	*/õn/
	*/ot/	*/ot/
	*/õt/	*/õt/
Cue constraints	*[ot] /ot/	*[ot] /ot/
	*[ot] /õt/	*[ot] /õt/
	*[õt] /ot/	*[õt] /ot/
	*[õt] /õt/	*[õt] /õt/
	*[on] /on/	*[on] /on/
	*[on] /õn/	*[on] /õn/
	*[õn] /on/	*[õn] /on/
	*[õn] /õn/	*[õn] /õn/

Constraints for non-nasal context forms: **Constraints for nasal context forms:**
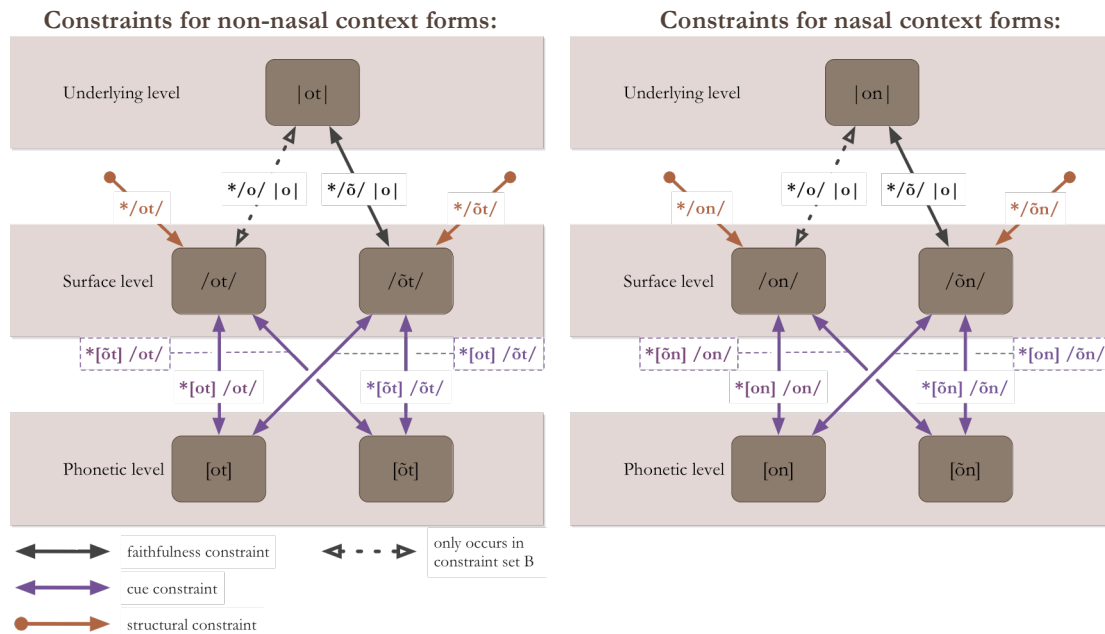
Figure 12: Schematic overview of all constraints in constraint set A and B

2.1.3   Initial ranking value

At the start of the learning phase, learners begin with an initial state of their grammar. In this initial state the grammar consists of one of the constraint sets described in Section 2.1.2, and the initial ranking values of those constraints. As a default, we have set all initial ranking values at an equal value of 100. However, in order to test the influence of initial ranking on the chosen surface forms (see research question 5 in Section 1.7) we have also created initial rankings of 101-110 for the cue constraint *[ot]/õt/, with all other constraints still ranking at 100. By giving this constraint a higher initial ranking, we expect the output path [ot]/õt/|ot| to become a less likely output. Figure 13 shows all possible perception output trees with [ot] and [õn] as input. All of these trees could occur, but we hypothesise that if the *[ot]/õt/ constraint is ranked higher than the others, Figure 13b and Figure 13c become less likely candidates, leaving as a possibility the trees in Figure 13a and Figure 13d.

a. branching:
[ot]/ot/|ot| &
[õn]/õn/|on|

b. branching:
[ot]/õt/|ot| &
[õn]/on/|on|

c. non-branching:
[ot]/õt/|ot| &
[õn]/õn/|on|

d. non-branching:
[ot]/ot/|ot| &
[õn]/on/|on|
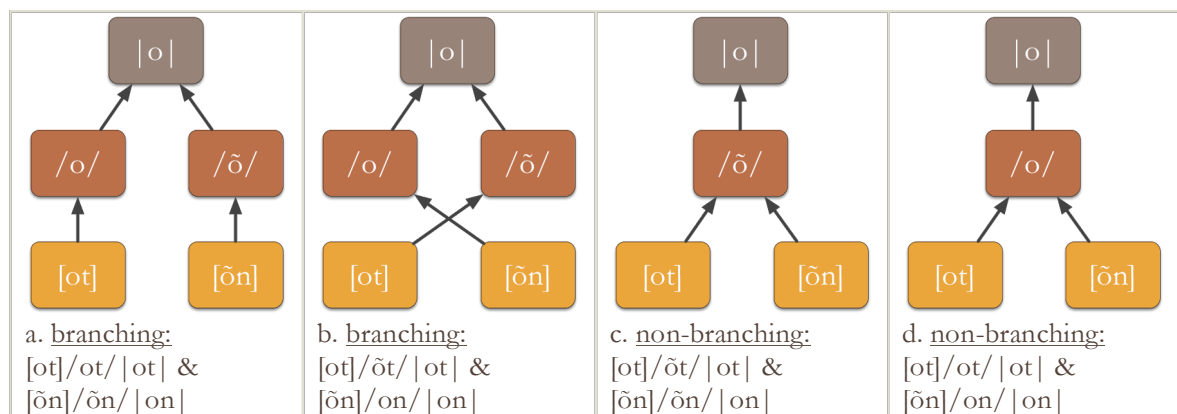
Figure 13: Possible perception output trees for [ot] & [õn] (consonants have been left out of the surface and underlying forms in the figure due to readability)

2.2 Testing phase

During the testing phase, each learner is asked to generate 10 optimal output paths for every possible input for perception and production. These perception inputs are [ot], [õt], [on] and [õn], and the production inputs are |ot| and |on|. If a learner generates 10 identical output paths for a certain input, his output for that input is labelled as consistent. If the learner produces different output paths, his output for that input is labelled as inconsistent. As is shown in Table 10 and 11, for each perception input, two different output paths are possible, while for each production input, four different output paths are possible.

Table 10: All possible perception outputs

perception input	possible outputs	
[ot]	[ot] /ot/ |ot|	[ot] /õt/ |ot|
[õt]	[õt] /ot/ |ot|	[õt] /õt/ |ot|
[on]	[on] /on/ |on|	[on] /õn/ |on|
[õn]	[õn] /on/ |on|	[õn] /õn/ |on|

Table 11: All possible production outputs

production input	possible outputs			
|ot|	[ot] /ot/ |ot|	[ot] /õt/ |ot|	[õt] /ot/ |ot|	[õt] /õt/ |ot|
|on|	[on] /on/ |on|	[on] /õn/ |on|	[õn] /on/ |on|	[õn] /õn/ |on|

3 Results

Experiments were run with the different configurations in variables (constraint sets, see Section 2.1.2; input distributions, see Section 2.1.1.1, initial constraint rankings, see Section 2.1.3), and then output results from the testing phase (see Section 1.7) were compiled and compared. In the following paragraphs we discuss these results.

3.1 Constraint sets

3.1.1 Perception

First, we discuss the perception results for the two constraint sets A and B. As is described in Section 2.1.2, constraint set A has asymmetric faithfulness constraints, while constraint set B has symmetric faithfulness constraints (see Table 9 for a full list of the constraints in both sets). Figure 14 shows the output paths chosen in perception by 1,000 learners in input distribution scenarios 1, 2 and 3 (see Table 8 and Figure 11) and with an asymmetric or a symmetric constraint set. After completing the learning process, these learners were asked to generate perception output paths for the phonetic form inputs [ot], [õt], [on] and [õn] (see Table 10). If a learner generates different output paths for one perception input, his output for that input category has been marked as

	Constraint set A - Asymmetric			Constraint set B - Symmetric		
	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
■ one or more outputs inconsistent	1000	1000	1000	12	4	7
■ [ot]/õt/ & [õt]/õt/ & [on]/on/ & [õn]/on/	0	0	0	506	510	503
■ [ot]/ot/ & [õt]/ot/ & [on]/õn/ & [õn]/õn/	0	0	0	482	486	490
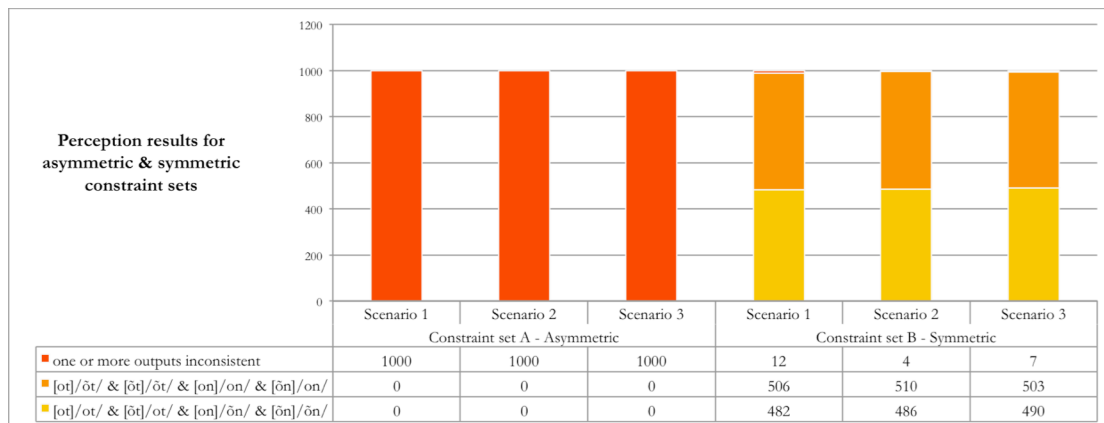
Figure 14: Results for constraint sets A and B – Perception

inconsistent. As we can see in Figure 14, none of the learners with constraint set A converge on consistent output for all of the input categories, regardless of the input distribution scenario. In the Appendix, a more detailed version of Figure 14 is included that shows exactly for which inputs learners generated inconsistent outputs. This Appendix Figure 1 shows that for 90% of learners with the constraint set A, outputs for all four input categories were inconsistent. Learners with constraint set B, on the other hand, do generate consistent outputs: 98.8% of learners in scenario 1 generated consistent outputs in all perception input categories, to 99.6% of learners in scenario 2, and 99.3% of learners in scenario 3.

Of those learners that do generate consistent output in all perception categories with constraint set B (calculated over all input distribution scenarios, see Figure 15), 49% of learners generate branching output paths with non-nasal vowel surface forms in non-nasal contexts (i.e. /ot/) and nasal surface forms in nasal contexts (i.e. /õn/), regardless of the nasalisation of the vowel in the phonetic form. The other 51% generate branching output paths with nasal vowel surface forms in non-nasal contexts (i.e. /õt/) and non-nasal vowel surface forms in nasal contexts (i.e. /on/).
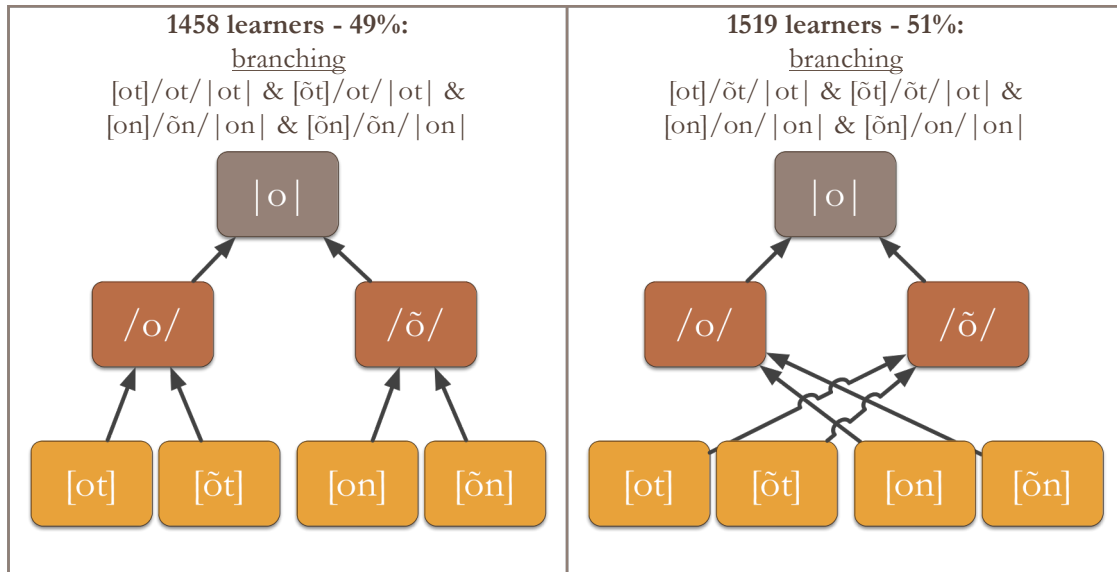
Figure 15: Perception results for consistent learners (total: 2977) of constraint set B

3.1.2 Production

Figure 16 shows the results of learners' production output paths when given as input the underlying forms |ot| and |on|. Whereas in perception there is a clear difference between results of learners with constraint set A and learners with constraint set B, this difference is not discernable in Figure 16. In learners of both constraint set A and constraint set B, learning with input Scenario 1 almost exclusively results in inconsistent output for one or more inputs, while in Scenario 2 and 3 a third of outputs are inconsistent, one third is branching with /ot/&/õn/ outputs and one third is branching with /õt/&/on/ outputs. However, the more detailed view of the inconsistent outputs in Appendix Figure 2 reveals a marked difference in inconsistent results for Scenario 1: 97.5% of learners with constraint set A generate inconsistent output for both |ot| and |on| inputs, while only 9.6% of learners with constraint set B do the same. The other 90.2% of learners in constraint set B only generate inconsistent outputs for |on|, while generating consistent output for |ot|. As such, learners with constraint set B produce results for scenario 1 that could be expected: input distribution scenario 1 provides little to no evidence in its language data for |on| (see Section 2.1.1.1), and consequently learners with this input distribution cannot generate consistent production output paths for that underlying form input. Learners with constraint set B do produce consistent production output for |ot| inputs in scenario 1, as those underlying forms are amply represented in the language learning data. As is shown in Appendix Figure 2, learners

	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
	Constraint set A - Asymmetric			Constraint set B - Symmetric		
■ one or more outputs inconsistent	1000	348	398	998	332	392
■ [ot]/õt/|ot| & [õn]/on/|on|	0	330	301	1	335	314
■ [ot]/ot/|ot| & [õn]/õn/|on|	0	322	301	1	333	294
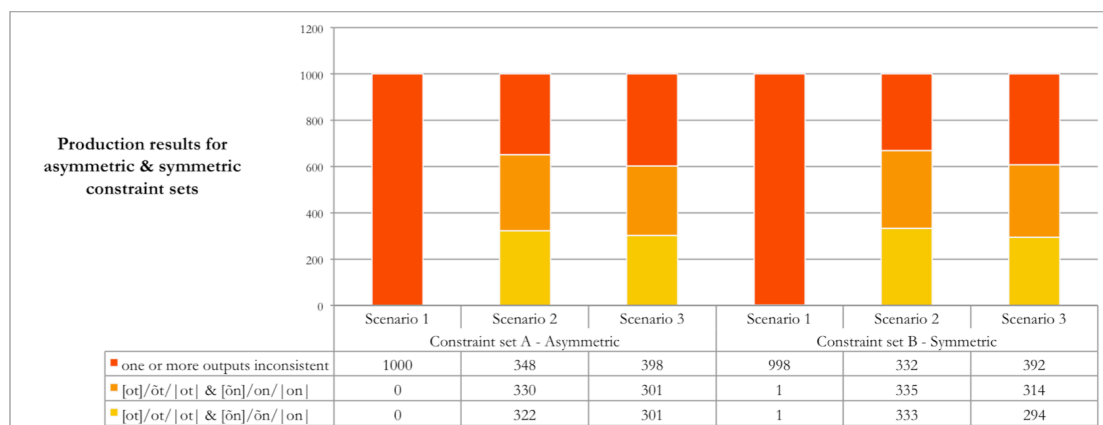
Figure 16: Results for constraint sets A and B - Production

with constraint set A do not produce consistent production output for either underlying form input in scenario 1. There is no discernable difference between the outputs of learners with constraint set A and B with input scenarios 2 and 3. The impact of the input distribution scenarios themselves on the results is further discussed in Section 3.3. Overall, it is evident that learners with constraint sets with symmetric faithfulness constraints generate significantly more consistent output paths than those with asymmetric constraint sets, in both perception and production. Therefore, in the following sections we compare only the results from learners with constraint set B. Another point of note is that none of the learners in perception or production have generated non-branching output paths such as those shown in Figure 13c or Figure 13d. This preference for branching output paths is further discussed in Section 4.3.

3.2 Initial constraint rankings

In order to find the answer to research question 5 in Section 1.7 on the effect of initial constraint ranking on the learners' choice of surface forms, the experiment was run on groups of 1,000 learners with different initial constraint rankings (100-110) only for the cue constraint *[ot]/õt/. This particular constraint was chosen because an output path like [ot]/õt/|ot| is just as plausible as an output path such as [ot]/ot/|ot| with a grammar that ranks all constraints equally. Since [ot]/ot/|ot| is theoretically the preferred output, there needs to be some reason inside the grammar in order for it to be chosen over [ot]/õt/|ot|. With this experiment we test the possibility that this reason is a higher initial ranking of *[ot]/õt/ in comparison to the other constraints. The expected result would be that the higher this constraint is ranked initially, fewer learners will choose /õt/ as a surface form in either perception or production. Results for these experiments are given in Figures 17-22.

3.2.1 Perception

Figures 17, 19 and 21 show the results for different initial constraint rankings in perception, for input distribution scenarios 1, 2 and 3 respectively. These figures show that there is no great difference in the perception results for the different input distribution scenarios; the major trend is the same for all three figures. The left-most column in the figures represents the perception outputs for learners with a grammar that has no initial constraint ranking for any constraints (i.e. the constraint *[ot]/õt/ is ranked at 100, just like all the other constraints). This column shows that when there is no initial constraint ranking, most learners generate either of the following combined perception outputs, in a near equal distribution (see also Figure 15):

1. [ot]/ot/|ot| & [õt]/ot/|ot| & [on]/õn/|on| & [õn]/õn/|on|
2. [ot]/õt/|ot| & [õt]/õt/|ot| & [on]/on/|on| & [õn]/on/|on|

Figures 17, 19 and 21 show that as the initial ranking value of *[ot]/õt/ gets greater, the proportion of learners who choose that second option gets smaller. From initial ranking value 106 on, there are no learners left who choose these /õt/&/on/ surface forms.

3.2.2 Production

Figures 18, 20 and 22 show production results for the different initial constraint ranking settings for leaners with input scenarios 1, 2 and 3. We look first at Figure 18, which describes learners' production outputs with input distribution scenario 1. Just like in the perception figures, the leftmost column shows the learners' production output when all constraints are initially (before learning) ranked at 100. In this situation, with no

Perception results for initial ranking values 100-110 (input distribution scenario 1)

	100	101	102	103	104	105	106	107	108	109	110
[ot]/õt/ & [õt]/ot/ & inconsistent & inconsistent	0	0	0	0	0	0	0	0	0	0	1
[ot]/õt/ & [õt]/õt/ & [on]/on/ & inconsistent	3	2	0	1	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/õn/ & inconsistent	5	6	2	7	8	9	5	5	4	3	10
[ot]/õt/ & [õt]/õt/ & inconsistent & [õn]/on/	4	3	1	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & inconsistent & [õn]/õn/	3	1	8	5	4	5	8	3	5	2	4
[ot]/õt/ & inconsistent & [on]/õn/ & [õn]/õn/	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & inconsistent & [on]/on/ & [õn]/õn/	0	0	0	0	0	0	1	0	0	0	0
inconsistent & [õt]/ot/ & [on]/õn/ & [õn]/õn/	2	3	2	3	1	3	2	1	1	0	0
inconsistent & [õt]/ot/ & [on]/on/ & [õn]/õn/	2	1	0	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/on/ & [õn]/õn/	495	283	146	49	8	1	0	0	0	0	0
[ot]/õt/ & [õt]/ot/ & [on]/õn/ & [õn]/õn/	486	701	841	935	979	982	984	991	990	995	985

Figure 17: Initial ranking values 100-110 - Scenario 1 - Perception

Production results for initial ranking values 100-110 (input distribution scenario 1)

	100	101	102	103	104	105	106	107	108	109	110
inconsistent & inconsistent	86	87	99	105	91	99	100	110	99	91	104
inconsistent & [õn]/on/ & |on|	0	0	0	0	0	1	0	1	0	0	0
inconsistent & [õn]/on/ & |on|	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & |ot| & inconsistent	457	261	134	46	8	1	0	0	0	0	0
[ot]/õt/ & |ot| & inconsistent	454	650	767	845	900	898	900	889	898	909	894
[ot]/õt/ & |ot| & [õn]/on/ & |on|	1	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & |ot| & [on]/on|	1	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & |ot| & [õn]/|on|	1	0	0	2	0	0	0	0	1	0	1
[ot]/õt/ & |ot| & [on]/|on|	0	2	0	2	1	1	0	0	2	0	1

Figure 18: Initial ranking values 100-110 - Scenario 1 - Production

Perception results for initial ranking values 100-110 (input distribution scenario 2)

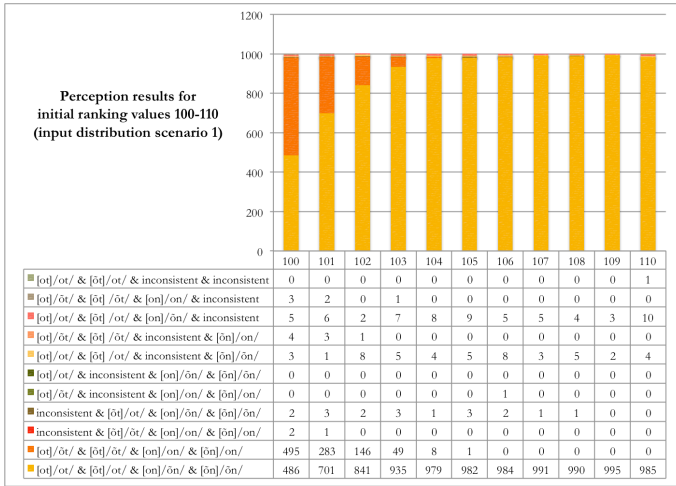|	100	101	102	103	104	105	106	107	108	109	110
[ot]/õt/ & [õt]/ot/ & inconsistent & inconsistent	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/on/ & inconsistent	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/õn/ & inconsistent	3	1	4	3	1	4	2	4	3	3	1
[ot]/õt/ & [õt]/õt/ & inconsistent & [õn]/on/	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/on/ & [õn]/õn/	0	0	1	1	1	1	0	0	0	0	0
[ot]/õt/ & inconsistent & [on]/on/ & [õn]/õn/	1	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & inconsistent & [on]/on/ & [õn]/õn/	0	0	0	0	0	0	0	0	0	0	0
inconsistent & [õt]/ot/ & [on]/õn/ & [õn]/õn/	1	1	3	2	1	1	1	0	1	0	0
inconsistent & [õt]/ot/ & [on]/on/ & [õn]/õn/	2	1	1	1	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/on/ & [õn]/on/	499	253	100	25	4	1	0	0	0	0	0
[ot]/õt/ & [õt]/ot/ & [on]/õn/ & [õn]/õn/	494	744	891	968	993	993	997	996	996	997	999

Figure 19: Initial ranking values 100-110 - Scenario 2 – Perception

Production results for initial ranking values 100-110 (input distribution scenario 2)

	100	101	102	103	104	105	106	107	108	109	110
inconsistent & inconsistent	31	29	30	34	29	28	33	28	46	26	33
inconsistent & [õn]/õn/ & |on|	76	106	135	136	132	129	143	132	140	146	147
inconsistent & [õn]/on/ & |on|	60	37	16	5	0	0	0	0	0	0	0
[ot]/õt/ & |ot| & inconsistent	82	47	18	2	0	0	0	0	0	0	0
[ot]/õt/ & |ot| & inconsistent	63	116	137	147	192	161	152	163	149	145	157
[ot]/õt/ & |ot| & [õn]/on/ & |on|	340	163	60	16	4	1	0	0	0	0	0
[ot]/õt/ & |ot| & [on]/|on|	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & |ot| & [õn]/õn/ & |on|	348	502	604	660	643	681	672	677	665	683	663
[ot]/õt/ & |ot| & [on]/õn/ & |on|	0	0	0	0	0	0	0	0	0	0	0

Figure 20: Initial ranking values 100-110 - Scenario 2 - Production

Perception results for initial ranking values 100-110 (input distribution scenario 3)

	100	101	102	103	104	105	106	107	108	109	110
[ot]/õt/ & [õt]/ot/ & inconsistent & inconsistent	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/on/ & inconsistent	3	2	1	1	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/õn/ & inconsistent	1	2	3	1	2	2	2	0	1	2	1
[ot]/õt/ & [õt]/õt/ & inconsistent & [õn]/on/	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & inconsistent & [õn]/õn/	0	0	2	1	1	1	0	0	1	1	0
[ot]/õt/ & inconsistent & [on]/õn/ & [õn]/õn/	1	2	0	0	0	0	0	0	0	0	0
[ot]/õt/ & inconsistent & [on]/on/ & [õn]/õn/	0	0	0	0	0	0	0	0	0	0	0
inconsistent & [õt]/ot/ & [on]/õn/ & [õn]/õn/	1	1	2	5	0	6	3	0	1	0	0
inconsistent & [õt]/ot/ & [on]/on/ & [õn]/õn/	3	0	1	0	0	0	0	0	0	0	0
[ot]/õt/ & [õt]/õt/ & [on]/on/ & [õn]/on/	520	275	95	30	4	1	0	0	0	0	0
[ot]/õt/ & [õt]/ot/ & [on]/õn/ & [õn]/õn/	471	718	896	962	993	990	995	1000	997	997	999

Figure 21: Initial ranking values 100-110 - Scenario 3 - Perception

Production results for initial ranking values 100-110 (input distribution scenario 3)

	100	101	102	103	104	105	106	107	108	109	110
inconsistent & inconsistent	29	43	33	36	34	42	34	39	34	42	35
inconsistent & [õn]/õn/ & |on|	58	62	84	99	99	101	89	103	88	95	101
inconsistent & [õn]/on/ & |on|	42	23	14	2	3	0	0	0	0	0	0
[ot]/õt/ & |ot| & inconsistent	142	65	22	4	1	0	0	0	0	0	0
[ot]/õt/ & |ot| & inconsistent	122	195	231	222	231	241	257	251	238	244	265
[ot]/õt/ & |ot| & [õn]/on/ & |on|	327	180	56	25	0	1	0	0	0	0	0
[ot]/õt/ & |ot| & [on]/|on|	0	0	0	0	0	0	0	0	0	0	0
[ot]/õt/ & |ot| & [õn]/õn/ & |on|	280	432	560	612	632	615	620	607	640	619	599
[ot]/õt/ & |ot| & [on]/õn/ & |on|	0	0	0	0	0	0	0	0	0	0	0

Figure 22: Initial ranking values 100-110 - Scenario 3 - Production

additional initial constraint ranking for the constraint *[ot]/õt/, we can see that almost all learners (997/1000) with input distribution scenario 1 are not able to generate consistent production output for |on|. As was shown in Section 2.1.1, inputs with |on| make up only 1% of language data input in scenario 1, so learners with this input distribution have received very little evidence for these cases during the learning phase. However, most learners do generate consistent production output for |ot| (914/1000), which makes up

97% of language input in scenario 1. Of all 1,000 learners in this leftmost column, 454 learners generated [ot]/ot/|ot| for |ot| and inconsistent outputs for |on|, whereas 457 learners generated [ot]/õt/|ot| for |ot| and inconsistent outputs for |on|. As the initial constraint ranking of *[ot]/õt/ increases, the number of learners who generate [ot]/õt/|ot| for |ot| diminishes, until at initial constraint ranking 105 only one learner with that output is left, and none are left by initial constraint ranking 106. With initial constraint ranking 105-110 and input scenario 1, the number of learners per category is more or less stable, with around 90% of learners generating [ot]/ot/|ot| for |ot| and inconsistent outputs for |on|, and around 10% generating inconsistent outputs for both |ot| and |on|.

Figures 20 and 22 show the production results for learners with input distribution scenario 2 (49% [ot]|ot|; 49% [õn]|on|) and scenario 3 (69% [ot]|ot|; 29% [õn]|on|). We discuss the differences between these two figures in the following Section 3.3. For now we limit our discussion to the overall effect of the initial constraint rankings common to the two figures. In the situation without any initial constraint ranking, we can see that in these input distributions, around 30% of learners generate [ot]/ot/|ot| for |ot| and [õn]/õn/|on| for |on|. Another approximate 30% of learners generate [ot]/õt/|ot| for |ot| and [õn]/on/|on| for |on|, and the remaining 40% of learners generate inconsistent outputs for either or both inputs. This 60%-40% consistent to inconsistent learners ratio remains approximately the same for all initial constraint rankings. However, there is a change in the inner composition of these groups. Where consistent learners are divided between /ot/&/õn/ and /õt/&/on/ surface form outputs with no initial constraint ranking, from initial constraint ranking 106 onward all consistent learners choose /ot/&/õn/ as surface form outputs. Similarly, learners who generate inconsistent outputs for one category no longer generate /õt/ or /on/ as surface form for their consistent category from initial constraint ranking 105 onward. Overall, we can conclude that higher initial constraint rankings for *[ot]/õt/ have the effect that learners in both perception and production generate fewer output paths with the surface forms /õt/ and /on/, favouring instead surface forms /ot/ and /õn/.

3.3 Input distributions

Figures 23 and 24 show the output paths for both perception and production chosen by at least 10% of learners in each input distribution scenario, see also Appendix Table 1. Figure 23 shows the output paths for learners who started with an initial constraint ranking of 100 for all constraints, and Figure 24 shows them for learners who start with an initial ranking of 110 for the constraint *[ot]/õt/ and 100 for all other constraints. The percentages shown were computed from the tables in Appendix Table 2. Perception paths are indicated with grey arrows pointing in the perception direction, while paths that are generated in perception and production are indicated with black arrows pointing in both directions. In some cases no consistent production output was generated for one or both of the production inputs; in those cases no production paths are shown for the inconsistent outputs.

In the paragraphs below, we explain the graphs shown in Figure 23. A summary of these results is available in Table 12.
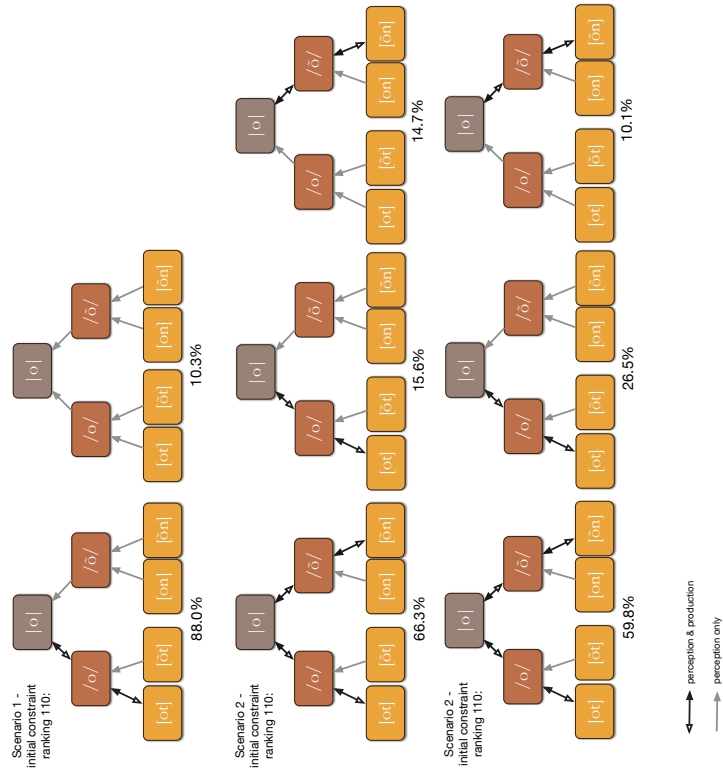
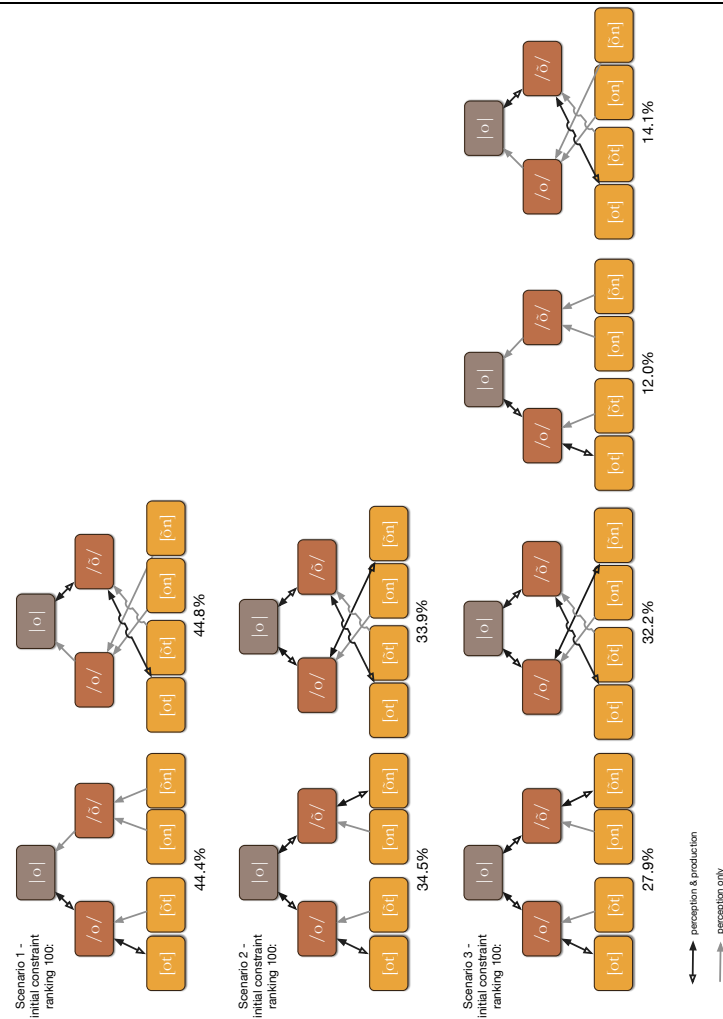**Figure 24: Perception & production output paths for learners with initial constraint ranking 110 for constraint *[ot]/õt/**

Figure 23: Perception & production output paths for learners with all constraints initially ranked at 100.

3.3.1 With initial constraint ranking 100

Figure 23 shows that for learners with input distribution scenario 1 (this scenario consisted of 97% [ot]|ot| as input, see Section 2.1.1.1) that had an initial grammar where all constraints were ranked at 100, the largest two groups of learners both give branching output paths:

1-a The first group (44.4%) of learners generates a branching perception path with non-nasal context phonetic forms (i.e. phonetic forms ending in [t]) being perceived as non-nasal surface forms (/ot/) and nasal context phonetic forms (i.e. phonetic forms ending in [n]) being perceived as nasalised surface forms (/õn/). In production, this group of learners generates consistent output for |ot| (namely the path [ot]/ot/|ot|), but no consistent output is generated for |on|.

1-b The other group (44.8%) of learners with input distribution scenario 1 also generates branching perception output paths, but now with non-nasal context phonetic forms being perceived as nasalised surface forms (/õt/) while nasal context phonetic forms are perceived as non-nasal surface forms (i.e. /on/). Again, this group does generate consistent output for |ot| (now [ot]/õt/|ot|), but no consistent output is generated for |on|.

Most learners with input distribution scenario 2 also fall into two major groups:

2-a Like with the Scenario 1 learners, one group (34.5%) of learners generates perception paths where non-nasal context phonetic forms are perceived as non-nasal surface forms (/ot/) and nasal context phonetic forms are perceived as nasalised surface forms (/õn/). However, unlike scenario 1 learners, most scenario 2 learners generate consistent output paths for both production inputs. This first group generates [ot]/ot/|ot| for |ot|, and [õn]/õn/|on| for |on|.

2-b The other group (33.9%) generates the same perception and production paths, only with the nasalised and non-nasalised surface forms switched around: the perception surface form is nasalised for non-nasal contexts (/õt/) and not nasalised in nasal contexts (/on/) and while phonetic forms in production are the same as in the first group, again the surface forms have switched, resulting in the production outputs [ot]/õt/|ot| for |ot| and [õn]/on/|on| for |on|.

Learners who were assigned input distribution scenario 3 can be divided into four groups:

3-a The first group (27.9%) has the same results as described above for group 2-a;

3-b The second group (32.2%) has the same results as described above for group 2-b;

3-c The third group (12.0%) has the same results as described above for group 1-a;

3-d The fourth group (14.1%) has the same results as described above for group 1-b.

Now that we have described the different groups of learners with different results for these input scenarios with default (100) initial constraint ranking, the following points stand out as the most prominent aspects of these results:

- In all three scenarios, learners produce branching output paths in both perception and production. None of the learners in any of the scenarios produce non-branching output paths.
- The majority of all learners succeed in generating consistent perception output paths (Appendix Table 2 shows that only 1.1% of learners generates inconsistent perception output paths for any of the four inputs).
- In these three scenarios, learners choose either of the following perception paths (where V stands for the vowel regardless of nasalisation):
  - [Vt]/ot/|ot| and [Vn]/õn/|on|

- o [Vt]/õt/|ot| and [Vn]/on/|on|
- If learners generate consistent production outputs for both |ot| and |on|, they choose either of the following production paths:
    - o [ot]/ot/|ot| and [õn]/õn/|on|
    - o [ot]/õt/|ot| and [õn]/on/|on|
- If learners only produce consistent production outputs for |ot|, they generate either of the following production paths:
    - o [ot]/ot/|ot|
    - o [ot]/õt/|ot|
- Learners with input distribution scenario 1 are only able to generate consistent production outputs for |ot|. They are unable to generate consistent production outputs for |on|.
- The majority of learners with input distribution scenario 2 generates consistent production outputs for both |ot| and |on|.
- Most learners with input distribution scenario 3 generate consistent production outputs for both |ot| and |on|, but a significant percentage (over 25%) is unable to generate consistent production outputs for |on|.

3.3.2   With initial constraint ranking 110

Figure 24 shows the results for learners with different input distribution scenarios who have all been assigned an initial constraint ranking of 110 for the constraint *[ot]/õt/, all other constraints are initially ranked at the default 100. For every input distribution scenario, the results that make up 10% or more of learners are shown in Figure 24. For the complete table of results, see Appendix Table 2.

Learners with input distribution scenario 1 can be divided into two groups:

1-a   88% of learners with input distribution scenario 1 generate perception paths with non-nasalised vowel surface forms for non-nasal context phonetic forms (/ot/) and nasalised vowel surface forms for nasal context phonetic forms (/õn/), and generate production paths with inconsistent output for |on| and consistent output for |ot| (namely [ot]/ot/|ot|).

1-b   10.3% of learners with this scenario produce the same perception output paths as group 1-a, but are unable to generate consistent production output for both |ot| and |on| inputs.

Learners with input distribution scenario 2 can be divided into three groups:

2-a   Most learners (66.3%) with this input distribution generate consistent perception and production output paths for all inputs. Perception paths are again non-nasalised vowel surface forms for non-nasal context phonetic forms (/ot/) and nasalised vowel surface forms for nasal context phonetic forms (/õn/), and production paths are [ot]/ot/|ot| and [õn]/õn/|on|.

2-b   However, 15.6% of learners in this category have the same perception paths as group 2-a but are unable to generate consistent production output for |on|. Production outputs for |ot| are consistent with those of group 2-a: [ot]/ot/|ot|.

2-c   Another 14.7% of learners with input distribution scenario 2 generate the same perception paths as 2-a, but are unable to generate consistent production output for |ot|. Production outputs for |on| are again consistent with those in group 2-a: [õn]/õn/|on|.

Learners with input distribution scenario 3 can be divided into the same three groups as those with input distribution scenario 2, but the groups differ in size:

3-a   Like in scenario 2, most learners in scenario 3 (here 59.8%) generate consistent perception and production paths. Learners in this group generate the same output paths as those described in 2-a.

Table 12: Summary of results for learners with different input distribution scenarios

	Initial ranking value 100	Initial ranking value 110
Scenario 1 *Inputs:* 97% [ot] |ot| 1% [õt] |ot| 1% [on] |on| 1% [õn] |on|	- branching output paths - *perception* - 2 possibilities: either {[Vt]/ot/ |ot| & [Vn]/õn/ |on|} or {[Vt]/õt/ |ot| & [Vn]/on/ |on|} - *production* - no consistent output for |on|; 2 possibilities for |ot|	- branching output paths - *perception* - 1 possibility:{[Vt]/ot/ |ot| & [Vn]/õn/ |on|} - *production* - no consistent output for |on|; 1 possibility for |ot|; at times no consistent output for |ot| either
Scenario 2 *Inputs:* 49% [ot] |ot| 1% [õt] |ot| 1% [on] |on| 49% [õn] |on|	- branching output paths - *perception* - 2 possibilities: either {[Vt]/ot/ |ot| & [Vn]/õn/ |on|} or {[Vt]/õt/ |ot| & [Vn]/on/ |on|} - *production* - consistent output for both |ot| and |on|; 2 possibilities: either {[ot]/ot/ |ot| & [õn]/õn/ |on|} or {[ot]/õt/ |ot| & [õn]/on/ |on|}	- branching output paths - *perception* - 1 possibility: {[Vt]/ot/ |ot| & [Vn]/õn/ |on|} - *production*: mostly consistent output for both |ot| and |on|; 1 possibility: {[ot]/ot/ |ot| & [õn]/õn/ |on|}; significant percentage (30%) no consistent output for either |ot| or |on|
Scenario 3 *Inputs:* 69% [ot] |ot| 1% [õt] |ot| 1% [on] |on| 29% [õn] |on|	- branching output paths - *perception*: 2 possibilities - either {[Vt]/ot/ |ot| & [Vn]/õn/ |on|} or {[Vt]/õt/ |ot| & [Vn]/on/ |on|} - *production*: mix of scenarios 1 & 2: around 60% consistent output for both |ot| and |on|; more than 25% only consistent output for |ot|, not for |on|	- branching output paths - *perception* - 1 possibility: {[Vt]/ot/ |ot| & [Vn]/õn/ |on|} - *production*: mostly consistent output for both |ot| and |on|; 1 possibility: {[ot]/ot/ |ot| & [õn]/õn/ |on|}; significant percentage (>35%) no consistent output for either |ot| or |on|

3-b Over a quarter of learners with scenario 3 (26.5%) generate the same output paths as described above in 2-b (i.e. no consistent production output for \|on\|).

3-c A little more than 10% of learners with scenario 3 generate the same output paths as described above in 2-c (i.e. no consistent production output for \|ot\|).

We can now summarise the results listed above as such:

- All learners (regardless of input distribution scenario or initial constraint ranking) generate branching output paths in production and perception;
- Almost all learners (regardless of input distribution scenario or initial constraint ranking) generate consistent perception output paths;
- Learners with initial constraint ranking 110 for the constraint *[ot]/õt/ do not generate perception or production output paths with /õt/ as a surface form, whereas learners with default initial constraint ranking do;
- Learners with input distribution scenario 1 do not generate consistent output results for \|on\|, regardless of initial constraint ranking;
- Learners with input distribution scenario 2 generate the most consistent output results for production in comparison to the other scenarios;
- Learners with input distribution scenario 3 generate much more consistent output results for production than learners with scenario 1, but are still significantly less successful in generating consistent production outputs for \|on\| than learners with scenario 2.

4 Discussion

Having presented the results of our experiments above in Section 3, we now turn to the implications of these results on the issues and questions posed in Section 1, and in particular the research questions listed in Section 1.7.

4.1 (A)symmetric faithfulness constraints

Q1. Is the learners' assignment of surface forms influenced by the presence/absence of symmetrical faithfulness constraints between the surface form and underlying form?

The first research question posed in Section 1.7, and repeated here, concerns the effect of symmetric and asymmetric faithfulness constraints on the learners' assignment of surface forms. With asymmetric faithfulness constraints, we mean here the traditional constraint construction of only having a faithfulness constraint that militates against the connection between an underlying form that is different (unfaithful) to the linked surface form. A symmetric faithfulness construction has both this 'normal' faithfulness constraint, and the opposite constraint in the lines of "do not link this underlying form with a faithful surface form". The effects of these two faithfulness constraint constructions were tested by running simulations on 1,000 learners with two different constraint sets: one with asymmetric faithfulness constraints (constraint set A) and one with symmetric faithfulness constraints (constraint set B). For a more detailed description of the constraint sets used, see Section 2.1.2. The results of these simulations are presented in Section 3.1. Results showed that learners with constraint set A (asymmetric faithfulness constraints) are universally unable to generate consistent perception outputs, while the majority of learners with constraint set B (symmetric faithfulness constraints) did generate consistent perception output (see Figure 14). In production, there is a vast difference between learners with constraint set A and constraint set B who have been assigned input distribution scenario 1 (97% [ot]|ot| input): almost all learners with constraint set A produce inconsistent output for both |ot| and |on|, while most learners with constraint set B produce consistent output for |ot| but are unable to produce consistent output for |on| (see Appendix Figure 2). The difference in production between learners with the two constraint sets is much less pronounced in scenarios 2 and 3: in scenario 2, learners with constraint set B generate 1.6% more consistent production outputs than learners with constraint set A, and in scenario 3 this difference is only 0.6%.

To answer the question posed at the start of the paragraph: if learners make use of an asymmetric constraint set as opposed to a symmetric constraint set, there is a significant negative effect on learners' success in assigning a surface form in perception, as well as in production with input distribution scenario 1. In the other two input distribution scenarios, the difference in production results between learners with either constraint set is too small to conclude definitively that there is an effect.

Having proven that there is an effect, however, brings up another question: why can't learners generate consistent perception outputs (and some production outputs) with an asymmetric faithfulness constraint set?

We reproduced the learning simulation with a smaller set of candidates and constraints, by restricting them to forms with nasal context (see Appendix Table 3 for an overview of the altered constraint sets and candidate set), and with a small number of learners (n=10). With this smaller experiment we tested three different constraint sets: two asymmetric constraint sets and one symmetric constraint set. The asymmetric constraint sets were designed to have opposite faithfulness constraints (one had */on/|on|, and the other */õn/|on|) to test the direction of the asymmetric effect . In an effort to rule out any effects from the input distribution, we used three different input distributions, see Table 13.

Table 13: input distributions for (a)symmetric constraint set test

	[õn]|on|	[on]|on|
Distribution 1	99%	1%
Distribution 2	1%	99%
Distribution 3	50%	50%

All other settings, such as the number of inputs each learner received (800,000), were kept the same as in the main experiments described in Section 2. Results for these simplified experiments are shown in Appendix Table 4 and Appendix Table 5. Surprisingly, no difference was found in learners' ability to generate consistent output, since almost none of the learners generated consistent output. However, we did find that the difference in input distributions only had an influence in production, and then only on the phonetic form that was chosen. The symmetricity/asymmetricity of the constraint sets themselves had an influence on the surface forms chosen in both perception and production: learners with the asymmetric set with the constraint */õn/\|on\| were more likely to choose /õn/ surface forms in perception and production, while learners with the other asymmetric set (with the constraint */on/\|on\|) were more likely to generate perception and production paths with /on/ surface forms. Learners with the symmetric constraint set had no bias toward production or perception paths with either surface form. It is curious that in this test with a smaller candidate set and constraint set, no difference in consistency of outputs was found between asymmetric and symmetric constraint sets, when the difference is so pronounced in the main experiment (see Figure 14). A possible answer is that the reduction of the constraint set and candidate set in this test has led to a situation where there is too little contrast, or competition, for learning to succeed. Further investigation into the minimum number of constraints and candidates and a minimum amount of variation in the input needed for successful learning with partial inputs in OT may be helpful for further studies. However, this goes beyond the scope of the current thesis. To return to the constraints and candidates used in the main simulations of this thesis, we looked in detail at the constraint rankings for learners with the asymmetric and symmetric constraint sets described in Section 2.1.2. This was done in an attempt to discover why learners with the asymmetric constraint set are unable to produce consistent perception output while learners with the symmetric constraint set are able to do so. We simulated learning for one learner with constraint set A and one learner with constraint set B with 100,000 inputs in the same distribution as scenario 2 of the main experiment (i.e. 49% [ot]\|ot\|, 49% [õn]\|on\|, 1% [õt]\|ot\|, 1% [on]\|on\|). We now compare the resulting constraint rankings for both learners in Figure 25. Results in this figure show that the constraint rankings in the grammar of the learner with the asymmetric constraint set (A) have a far smaller range than those in the other learner's grammar. The learner with the symmetric constraint set (B) is able to re-rank the constraints in his grammar in such a way that there is enough distance between constraints that output is consistent, even when noise is added to the ranking value. Table 14 shows the outputs both learners generate when asked for 10 outputs for each possible input. The learner with constraint set A generates inconsistent output, while the learner with constraint set B generates consistent output for all input categories, even after only 100,000 inputs (in

Learner with constraint set A		Learner with constraint set B	
	ranking value		*ranking value*
* [on]/on/	103.142	*/õt/	110.150
* [õt]/ot/	103.135	*/on/	109.511
* [on]/õn/	102.543	* [õt]/õt/	107.938
* [õt]/õt/	102.443	* [on]/on/	107.573
*/ot/	100.940	* [ot]/õt/	102.212
*/on/	100.684	* [õn]/on/	101.938
*/õn/	99.316	*/+nas/|-nas|	100.639
*/õt/	99.060	*/-nas/|-nas|	99.361
*/+nas/|-nas|	98.376	* [on]/õn/	98.169
* [ot]/ot/	97.806	* [õt]/ot/	97.797
* [õn]/on/	97.541	* [õn]/õn/	92.320
* [õn]/õn/	96.774	* [ot]/ot/	92.052
* [ot]/õt/	96.616	*/õn/	90.489
		*/ot/	89.850

Figure 25: Constraint rankings of a learner with constraint set A and a learner with constraint set B after 100,000 inputs from input distribution scenario 2

Table 14: Outputs for learner with constraint set A and learner with constraint set B

input	Learner with constraint set A		Learner with constraint set B	
[ot]	[ot]/ot/|ot|	x5	[ot]/ot/|ot|	x10
	[ot]/õt/|ot|	x5		
[õt]	[õt]/ot/|ot|	x4	[õt]/ot/|ot|	x10
	[õt]/õt/|ot|	x6		
[on]	[on]/on/|on|	x2	[on]/õn/|on|	x10
	[on]/õn/|on|	x8		
[õn]	[õn]/on/|on|	x3	[õn]/õn/|on|	x10
	[õn]/õn/|on|	x7		
|ot|	[ot]/ot/|ot|	x2	[ot]/ot/|ot|	x10
	[ot]/õt/|ot|	x8		
|on|	[õn]/on/|on|	x3	[õn]/õn/|on|	x10
	[õn]/õn/|on|	x7		

the main simulations we presented all learners with 800,000 inputs). Closer examination of an excerpt of the first 100 learning steps for both learners (see Appendix Table 6 and Appendix Table 7) indicates that this lack of contrast in ranking values is most likely due to the fact that the learner with an asymmetric constraint set (A) does not re-rank constraints as often as the learner with a symmetric constraint set (B), and that when the learner with constraint set A re-ranked a constraint, he often later re-ranks that constraint back in the opposite direction (thus nullifying the first re-ranking). We can conclude that learners with a asymmetric constraint sets are unable to generate consistent output (at least with the number of constraints and candidate sets in our main simulations) because the asymmetry during learning results in a constraint ranking that does not provide enough contrast between constraints. An interesting avenue for further research is to explore the issue of learning with asymmetric constraint sets, to see whether this effect also applies in other types of learning, or if learners with these types of constraint sets only fail to generate consistent output under specific circumstances similar to ours.

4.2 Variation between learners

Q2. Do learners, given the same input distribution in the learning phase, assign the same surface forms to phonetic forms, or is there great variation between learners?

We found that successful learners with the same input distributions and default constraint ranking for all constraints are split almost equally into two groups when it comes to their perception output. One group (see Figure 26) generates perception paths that assign non-nasal surface forms to phonetic forms with non-nasal consonantal context (i.e. [ot]/ot/\|ot\| and [õt]/ot/\|ot\|). This group of learners also assigns nasal surface forms to phonetic forms with nasal consonantal context (i.e. [on]/õn/\|on\| and [õn]/õn/\|on\|). The other group of learners (see Figure 27) assigns the opposite surface forms, so nasal surface forms to phonetic forms with non-nasal context and non-nasal surface forms to phonetic forms with nasal context (i.e. [ot]/õt/\|ot\| and [õt]/õt/\|ot\| and [on]/on/\|on\| and [õn]/on/\|on\|). When we looked at the combination of perception and production (see Figure 23 and Figure 24), results showed that successful learners assign the same surface forms in production as they do in perception, and they

produce the phonetic forms that they have received the most evidence for in their input during learning. There are a couple of interesting things to note on this topic:
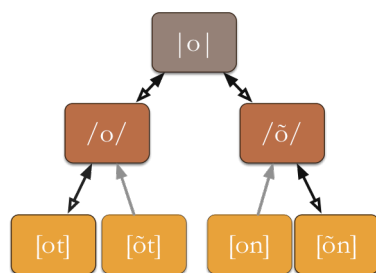
Figure 26: Branching surface form structure – [Vt]/ot/\|ot\| & [Vn]/õn/\|on\|
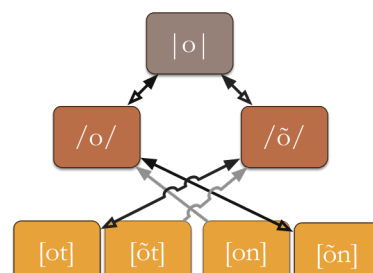
Figure 27: Branching surface form structure – [Vt]/õt/\|ot\| & [Vn]/on/\|on\|

- None of the learners assigned surface forms with the same vowel to all phonetic forms (so none of the learners assign non-branching output paths). This preference for branching surface structure is further discussed in Section 4.3.
- All of the learners assign nasal/non-nasal surface forms according to the consonantal context of the vowel in the phonetic form, not according to the nasal/non-nasal quality of the vowel itself (i.e. none of the learners formed branching paths such as the one depicted in Figure 28).

This second point, the non-occurrence of branching structures such as the one in Figure 28, may be due to some kind of economy in the OT learning mechanism. After all, the representation in Figure 28 contains a simplification of the surface and underlying forms: the actual paths in this structure are [ot]/ot/|ot|, [on]/on/|on|, [õt]/õt/|ot| and [õn]/õn/|on|. Note that this structure requires four different surface forms (/ot/, /on/, /õt/ and /õn/), whereas the perception paths that actually occurred most (see Figure 26 and Figure 27) require only two different surface forms (either /ot/ and /õn/ or /õt/ and /on/). The fact that none of the learners generated output paths in perception or production that required more than two surface forms suggests that one of the effects of the learning algorithm is to minimise the number of surface forms. The minimum number of surface forms in our simulations is two, since there is no possibility of having a single surface form in the current constraint set: there are no candidates that link phonetic forms with non-nasal consonantal context to surface forms with nasal consonantal context and vice versa, making paths like [ot]/on/|on| impossible.

What has led to the difference in the number of surface forms between the branching structures of Figure 26/Figure 27 and that of Figure 28, is the choice to include consonantal context in all levels of representation. After all, if there was no consonantal context in the surface level, there would be only two possible surface forms (/o/ and /õ/) instead of four. It might be that if information about context had not been included in the surface form, structures such as the one in Figure 28 would be learnable. In other words, the exclusion of possible perception paths could possibly be a consequence of our choice to include consonantal context in all levels of representation. However, we feel that to exclude such context from the abstract representations is to take a further step back from reality, where the majority of sounds is perceived and produced in context. Contextual information is undoubtedly of value to the nasalisation process itself, and could well be of use in assigning abstract structure.

Figure 28: Possible (non-occurring) branching perception path

4.3 Branching perception and production paths

Q3. Do learners favour one type of surface form assignment (i.e. branching vs. non-branching) over the other?

As was mentioned in the previous paragraph, the results of our experiments showed that none of the learners (regardless of constraint set, input distribution or initial constraint ranking) generated non-branching surface form structures in perception or production. All successful learners assigned a different surface form vowel to phonetic forms with nasal consonantal context than those with non-nasal consonantal context. In Section 4.2 we mentioned economy of surface forms as a possible explanation for the non-occurrence of a different type of branching surface form structure. However, this explanation does not account for the non-occurrence of non-branching structures such as the one pictured in Figure 29. The non-branching structure in Figure 29 represents a
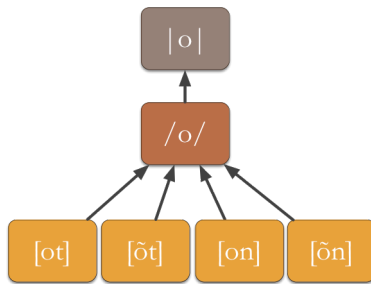
Figure 29: Non-occurring non-branching perception output path

learner who assigns all phonetic forms a surface form with a non-nasal vowel, and thus generates the following perception output paths: [ot]/ot/|ot|, [õt]/ot/|ot|, [on]/on/|on| and [õn]/on/|on|. This situation is one where there is variation on the phonetic level, but not on the surface (phonological) level. Note that for this structure, the same number of surface forms is required as are in the simulation's most occurring perception paths: here we would need /ot/ and /on/, while in the most occurring perception paths either /ot/ and /õn/ or /õt/ and /on/ are used. If minimization of the number of surface forms does not play a role here, since both options make use of the minimum of two surface forms, then what could be the cause for the learners' disuse of such a perception path? One might look for the answer in some sort of maximisation of distinctiveness: /ot/ and /õn/ are maximally distinctive in that each segment differs on the nasal dimension, whereas the surface forms /ot/ and /on/ only differ in one of the segments (i.e. the consonantal segment). The only problem is that this solution overlooks the fact that the OT decision mechanism does not make distinctions based on intrinsic qualities of the forms, it only takes notice of such differences if they are encoded in the constraint set. To the OT decision and learning mechanism, the difference between /ot/&/õn/ and /ot/&/on/ is the same as the difference between X&Y and X&Z. If the difference between these two sets of surface forms cannot be detected by the learning mechanism by looking at the internal structure of the surface forms, this difference must be encoded somehow in the constraint set. As such, a possible implication is that non-branching structures such as the one depicted in Figure 29 do not occur because they are not learnable with the current constraint set. The only constraints that have an influence on all output, regardless of the consonantal context, are the faithfulness constraints (*/õ/|o| and */o/|o|). Our tentative hypothesis is that, due to the surface form minimisation described in Section 4.2, two different surface forms will be utilised in the generation of output paths, and because of the way that the faithfulness constraints are formulated in the current constraint set, these two surface forms will contrast in the nasalisation of the vowel. We suggest that because this contrast in surface form vowels is encoded in the faithfulness constraints, and a minimum of two surface forms is needed in learning, the contrast is used to maximise the distinctiveness between the two surface form.

In future research this balance of maximisation of distinctiveness and minimisation of surface forms could be further investigated by including more phonetic forms in a similar experiment. By introducing more gradations of nasalisation and leaving open the option for additional surface forms in the constraint set, one could test whether learners always prefer two surface forms, or if the number of surface forms and the contrast in surface forms is influenced by the number of phonetic forms and faithfulness constraints.

4.4 Input distributions

Q4. Is the learners' assignment of surface forms influenced by the input distribution of phonetic forms in the learning phase?

As was shown in Section 3.3, the results of our simulations showed that the combined perception and production output of learners is different in each of our three input distribution scenarios (for now we disregard results for non-default initial constraint rankings, these will be discussed in Section 4.5). The most pronounced difference is that between scenarios 1 and the other two scenarios. This was to be expected, as the

difference in input distribution between scenario 1 and the others is also the most drastic, see Table 15.

Table 15: Input distribution scenarios

	[ot] |ot|	[õt] |ot|	[on] |on|	[õn] |on|
Scenario 1	97%	1%	1%	1%
Scenario 2	49%	1%	1%	49%
Scenario 3	69%	1%	1%	29%

Despite having received very little evidence of three out of four phonetic forms, learners with input distribution scenario 1 were overall successful in generating perception outputs for all phonetic forms (see Table 16 and Appendix Table 1). However, those learners are unable to generate consistent production output for the underlying form \|on\| (for which they have received very little evidence during learning). Learners who were assigned input distribution scenarios 2 and 3 during learning were more successful in generating consistent production outputs, as in both groups more than 60% of learners generated consistent production outputs. The main difference between learners with input distribution scenario 2 and those who were assigned input distribution scenario 3 during learning is that there are more learners who produce consistent production output with scenario 2 (68.8%, see Table 16) than with scenario 3 (60.7%). When we look more closely at the production outputs for these two groups, we can see that the skewness of the distribution in scenario 3 is reflected in the percentage of

Table 16: Consistency ratios per input distribution scenario for learners with default (100) initial constraint ranking for all constraints

	Learners with consistent perception output for all phonetic form inputs	Learners with consistent production output for |ot|	Learners with consistent production output for |on|	Learners with consistent production output for all underlying form inputs	Learners with consistent perception & production output for all inputs
Scenario 1	98.1%	91.4%	0.3%	0.3%	0.3%
Scenario 2	99.3%	83.3%	82.4%	68.8%	68.4%
Scenario 3	99.1%	87.1%	70.7%	60.7%	60.1%

learners who produce consistent output for \|ot\| and \|on\|: whereas learners with scenario 2 have comparable consistency results for \|ot\| and \|on\|, learners with scenario 3 generate significantly more consistent production outputs for \|ot\| (87.1%) than for \|on\| (70.7%), see Table 16. While both percentages are still quite high, we can see an effect of the altered input distribution in that the segment that is more frequently perceived during learning is more successfully (consistently) produced than a segment that is perceived relatively less frequently during learning.

To summarise, the input distribution learners are assigned during learning has a definite effect on their consistency in production outputs after learning, but learners generate consistent perception outputs irrespective of their input distribution. This discrepancy between perception and production is indicative of the power of the OT learning mechanism: even when presented with very little evidence of a phonetic form, learners are still able to assign consistent abstract structure to that input when it is perceived. However, when they are asked to produce phonetic output for an underlying form they have received little evidence for, they do not consistently produce the same phonetic output.

An avenue of further research would be to see how much evidence is generally needed for a learner to be able to consistently generate production paths with the underlying form as input. Is there a critical point of number of input pairs needed during learning, or is the percentage of this specific input in relation to the total number of input pairs important? If more phonetic forms are taken into account, an in-depth investigation of

the effect of the number of phonetic forms and more intricate input distributions is made possible.

4.5 Initial constraint ranking

Q5. In what way does the initial ranking of cue constraints influence the learners' assignment of surface forms to phonetic and underlying forms?

As can be seen by comparing Table 16 and Table 17, consistency results for learners with initial constraint ranking 100 or 110 for the constraint *[ot]/õt/ do not differ greatly. However, it does seem that learners with a higher initial constraint ranking for *[ot]/õt/ have slightly lower percentages of consistent production output, and slightly higher percentages of consistent perception output. The major difference between these two groups lies not in the consistency of the output, but in the choice of surface form. Learners with the higher initial constraint ranking do not produce the output path [ot]/õt/|ot|, which is quite common in the default initial ranking learners (see Appendix Table 1). Learners with the higher initial constraint ranking almost all choose /o/ surface forms for non-nasal context forms and /õ/ surface forms for nasal context forms. As was mentioned in Section 4.3, the OT learning mechanism does not take into account intrinsic value of surface forms; instead they are treated more as candidate labels while the actual decision mechanism is based on the constraints and their rankings. As such, while the learning mechanism without any special initial constraint rankings recognises that there are two categories (nasal context & non-nasal context) and assigns these two different surface forms, which of the surface forms is assigned to which category is equally distributed (see Figure 23). In order to arrive at a more 'phonologically logical' assignment of [ot] underlyingly having an /o/ vowel and [õn] having an underlyingly nasalised /õ/ vowel (instead of [ot] having an underlyingly nasalised vowel /õ/ while [õn] has an underlyingly non-nasalised vowel /o/), the dispreferred connection needs to be encoded in the constraint set. We have achieved this here by assigning the constraint *[ot]/õt/ a higher constraint ranking. As can be seen in Figures 17-22, this change in surface form assignment stabilises from an initial constraint ranking of around 106.

Table 17: Consistency ratios per input distribution scenario for learners with initial constraint ranking 110 for constraint *[ot]/õt/

	Learners with consistent perception output for all phonetic form inputs	Learners with consistent production output for |ot|	Learners with consistent production output for |on|	Learners with consistent production output for all underlying form inputs	Learners with consistent perception & production output
Scenario 1	98.5%	89.6%	0.2%	0.2%	0.2%
Scenario 2	99.9%	82.0%	81.0%	66.3%	66.3%
Scenario 3	99.9%	86.4%	70.0%	59.9%	59.8%

5 Conclusion

In the current thesis we have confirmed that, if presented with sufficient underlying-phonetic form pairs during learning, learners in a multidirectional GLA OT learning environment are able to generate perception and production paths including a surface form for any of those underlying forms or surface forms. We have somewhat elucidated when language evidence during learning is sufficient, finding that learners almost always succeed in finding some abstract structure to assign to phonetic forms in perception, but that for consistent production, 2% of 800,000 inputs (evidence for |on| in input distribution scenario 1) is insufficient. 30-70% of 800,000 inputs did prove to be sufficient evidence for the majority of learners. We have shown that in a situation with two different phonetic forms with different consonantal contexts, learners choose branching surface form structure over non-branching surface forms. Additionally, phonetic forms seem to be assigned surface forms according to their consonantal context, instead of the nasal quality of the vowel in the phonetic form. This raises questions about the maximisation of distinctiveness and the minimisation of the number of surface forms in an OT learning mechanism. We conjecture that these questions could be investigated by allowing more phonetic detail in further simulations by using more phonetic forms (such as a phonetic form with a semi-nasalised vowel and a phonetic form with a fully nasalised vowel), which also leads to a more extensive constraint set and more intricate input distributions. Another point of further investigation is the effect of asymmetric constraint sets on learning in a multidirectional GLA method. Our preliminary conclusion on this issue, based on results discussed in Sections 3.1 and 4.1, is that especially in perception having an asymmetric constraint set in their grammar negatively impacts learners. Learners with asymmetric constraint sets were shown not to be able to generate consistent perception outputs at all, while production seems to be mostly unaffected. However, in our attempt to reproduce this problem on a smaller scale (10 learners) with a smaller candidate set and constraint set (see Section 4.1), we did not see a clear difference in consistency of perception or production outputs between learners with a symmetric or an asymmetric constraint set. It seems there may be more factors at play with this asymmetricity of constraints, and as such we point it out as a future point of investigation.

6 References

Apoussidou, D. (2007). *The Learnability of Metrical Phonology*. Utrecht: LOT.

Boersma, P. (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the Universtity of Amsterdam* 21: 43-58.

Boersma, P. (2003). [Review of the book *Learnability in Optimality Theory*, by B. Tesar & P. Smolensky]. *Phonology* 20: 436-446.

Boersma, P. (2007). Some listener-oriented accounts of *h*-aspiré in French. *Lingua* 117: 1989-2054.

Boersma, P. (2011). A programme for bidirectional phonology and phonetics and their acquisition and evolution. In A. Benz & J. Mattausch (eds.): *Bidirectional Optimality Theory*, 33-72. Amsterdam: John Benjamins.

Boersma, P. & Hamann, S. (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217-270.

Boersma, P. & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Enquiry* 32: 45-86.

Botma, B. (2004). *Phonological aspects of nasality: An element-based dependency approach*. Dissertation. Amsterdam: University of Amsterdam.

Hajek, J. (2013). Vowel Nasalisation. In Dryer, M. S. & Haspelmath, M. (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Hamann, S. (2009). The learner of a perception grammar as a source of sound change. In Boersma, P. & Hamann, S. (eds.), *Phonology in Perception*. Berlin: Mouton de Gruyter. 111-149.

Kühnert, B. & Nolan, F. (1999). The origin of coarticulation. In Hardcastle, W.J. & Hewlett, N. (eds.), *Coarticulation – Theory, Data and Techniques*. Cambridge: Cambridge University Press. 7-30.

Prince, A. & Smolensky, P. (1993/2002). *Optimality Theory: Constraint Interaction in Generative Grammar*. [Revision of 1993 technical report, Rutgers University for Cognitive Science. Available on Rutgers Optimality Archive, ROA 537: version 8/2002].

Scheurup, L.C. (1973). *A cross-language study of vowel nasalization*. Master's Thesis. Columbus: Ohio State University.

Tesar, B.B. (1995). *Computational Optimality Theory*. Doctoral dissertation. Boulder: University of Colorado.

Tesar, B.B. & Smolensky, P. (1998). Learning Optimality-Theoretic grammars. *Lingua* 106: 161-196.

Young, L.H., Zajac, D.J., Mayo, R. & Hooper, C.R. (2001). Effects of Vowel Height and Vocal Intensity on Anticipatory Nasal Airflow in Individuals With Normal Speech. *Journal of Speech, Language, and Hearing Research* 44(1): 52-60.

Appendix

Perception results for asymmetric & symmetric constraint sets	Constraint set A - Asymmetric			Constraint set B - Symmetric		
	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
inconsistent & inconsistent & inconsistent & inconsistent	805	934	948	0	0	0
[ot]/õt/ & inconsistent & inconsistent & inconsistent	14	12	8	0	0	0
inconsistent & [õt]/õt/ & inconsistent & inconsistent	5	4	7	0	0	0
inconsistent & inconsistent & [on]/õn/ & inconsistent	25	10	6	0	0	0
inconsistent & inconsistent & inconsistent & [õn]/õn/	17	11	11	0	0	0
[ot]/õt/ & [õt]/õt/ & inconsistent & inconsistent	23	19	5	0	0	0
inconsistent inconsistent & [on]/õn/ & [õn]/õn/	111	10	15	0	0	0
[ot]/õt/ & [õt] /õt/ & [on]/on/ & inconsistent	0	0	0	2	2	2
[ot]/ot/ & [õt] /ot/ & [on]/õn/ & inconsistent	0	0	0	4	1	1
[ot]/õt/ & [õt] /õt/ & inconsistent & [õn]/on/	0	0	0	1	0	0
[ot]/ot/ & [õt] /ot/ & inconsistent & [õn]/õn/	0	0	0	3	0	1
[ot]/ot/ & inconsistent & [on]/õn/ & [õn]/õn/	0	0	0	0	0	1
[ot]/õt/ & inconsistent & [on]/on/ & [õn]/on/	0	0	0	0	0	1
inconsistent & [õt]/ot/ & [on]/õn/ & [õn]/õn/	0	0	0	2	0	0
inconsistent & [õt]/õt/ & [on]/on/ & [õn]/on/	0	0	0	0	1	1
[ot]/õt/ & [õt]/õt/ & [on]/on/ & [õn]/on/	0	0	0	506	510	503
[ot]/ot/ & [õt]/õt/ & [on]/õn/ & [õn]/õn/	0	0	0	482	486	490

Appendix Figure 1: Perception results for constraint set A and constraint set B – detailed view of inconsistent output results

Production output results for symmetric & asymmetric constraint sets	Constraint set A - Asymmetric			Constraint set B - Symmetric		
	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
inconsistent & inconsistent	971	42	41	96	31	56
inconsistent & [õn]/õn/|on|	1	71	53	0	72	51
inconsistent & [õn]/on/|on|	0	89	57	0	77	53
[ot]/õt/|ot| & inconsistent	27	68	124	456	89	113
[ot]/ot/|ot| & inconsistent	1	78	123	446	63	119
[ot]/õt/|ot| & [õn]/on/|on|	0	330	301	1	335	314
[ot]/ot/|ot| & [õn]/õn/|on|	0	322	301	1	333	294

Appendix Figure 2: Production results for constraint set A and constraint set B – detailed view of inconsistent output results

Appendix Table 1: Perception & production outputs for learners with all input distribution scenarios, and initial constraint ranking 100 or 110 for constraint *[ot]/ōt/

		Perception: [Vt]/ot/|ot| & [Vn]/ōn/|on| — Production: inconsistent output for |ot| and |on|	Perception: [Vt]/ot/|ot| & [Vn]/ōn/|on| — Production: |ot|/ot/|ot| & inconsistent for |on|	Perception: [Vt]/ot/|ot| & [Vn]/ōn/|on| — Production: inconsistent for |ot| & [ōn]/ōn/|on|	Perception: [Vt]/ōt/|ot| & [Vn]/on/|on| — Production: |ot|/ōt/|ot| & inconsistent for |on|
Initial ranking value 100	Scenario 1	4.1%	44.4%	0.0%	44.8%
	Scenario 2	1.2%	6.2%	7.5%	8.1%
	Scenario 3	1.4%	12.0%	5.8%	14.1%
Initial ranking value 110	Scenario 1	10.3%	88.0%	0.0%	0.0%
	Scenario 2	3.3%	15.6%	14.7%	0.0%
	Scenario 3	3.5%	26.5%	10.1%	0.0%

		Perception: [Vt]/ōt/|ot| & [Vn]/on/|on| — Production: |ot|/ōt/|ot| & [ōn]/on/|on|	Perception: [Vt]/ōt/|ot| & [Vn]/on/|on| — Production: |ot|/ōt/|ot| & [ōn]/on/|on|	Other
Initial ranking value 100	Scenario 1	0.1%	0.1%	6.5%
	Scenario 2	34.5%	33.9%	8.6%
	Scenario 3	27.9%	32.2%	6.6%
Initial ranking value 110	Scenario 1	0.1%	0.0%	1.6%
	Scenario 2	66.3%	0.0%	0.1%
	Scenario 3	59.8%	0.0%	0.1%

Appendix Table 2: Perception & production outputs for learners with all input distribution scenarios, and initial constraint ranking 100 or 110 for constraint *[ot]/õt/

Table (part 1)

Initial ranking value	Input distribution	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| & [õn]/õn/|on| [ot]/ot/|ot| [õn]/õn/|on|	inconsistent [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| [ot]/õt/|ot| [õn]/on/|on|	[ot]/õt/|ot| [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| [ot]/õt/|ot| [õn]/on/|on|	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| & [õn]/õn/|on| [ot]/ot/|ot| inconsistent	[ot]/õt/|ot| [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| [ot]/õt/|ot| inconsistent	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| & [õn]/õn/|on| [ot]/ot/|ot| inconsistent	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| & [õn]/õn/|on| inconsistent [õn]/õn/|on|	[ot]/õt/|ot| [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| inconsistent inconsistent	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| & [õn]/õn/|on| inconsistent inconsistent	[ot]/õt/|ot| [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| inconsistent inconsistent	[ot]/ot/|ot| inconsistent [on]/on/|on| & [õn]/on/|on| [ot]/ot/|ot| [õn]/õn/|on|	inconsistent [õt]/ot/|ot| [on]/õn/|on| & [õn]/õn/|on| [ot]/ot/|ot| [õn]/õn/|on|
100	Scenario 1	1	0	1	1	444	448	0	0	41	45	0	0
	Scenario 2	345	0	339	0	62	81	75	60	12	19	0	1
	Scenario 3	279	0	322	0	120	141	58	42	14	15	1	0
110	Scenario 1	1	1	0	0	880	0	0	0	103	0	0	0
	Scenario 2	663	0	0	0	156	0	147	0	33	0	0	0
	Scenario 3	598	0	0	0	265	0	101	0	35	0	0	0

Table (part 2)

Initial ranking value	Input distribution	inconsistent [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| [ot]/õt/|ot| inconsistent	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| & [õn]/õn/|on| inconsistent & inconsistent	[ot]/õt/|ot| [õt]/õt/|ot| [on]/on/|on| [õn]/on/|on| inconsistent inconsistent	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| inconsistent inconsistent	[ot]/õt/|ot| [õt]/õt/|ot| [on]/on/|on| & [õn]/õn/|on| inconsistent inconsistent	[ot]/ot/|ot| [õt]/ot/|ot| inconsistent [ot]/ot/|ot| inconsistent	[ot]/õt/|ot| [õt]/õt/|ot| inconsistent [ot]/õt/|ot| inconsistent	[ot]/ot/|ot| inconsistent [on]/õn/|on| & [õn]/õn/|on| [ot]/ot/|ot| inconsistent	inconsistent [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| [ot]/õt/|ot| inconsistent	[ot]/õt/|ot| [õt]/õt/|ot| [on]/on/|on| & [õn]/on/|on| [ot]/õt/|ot| inconsistent	[ot]/ot/|ot| [õt]/ot/|ot| [on]/õn/|on| inconsistent inconsistent
100	Scenario 1	0	0	0	5	3	3	4	0	2	2	0
	Scenario 2	1	2	2	1	0	0	0	1	0	1	0
	Scenario 3	3	0	2	1	1	0	0	0	1	0	0
110	Scenario 1	0	0	0	9	0	4	0	0	0	0	1
	Scenario 2	0	0	0	1	0	0	0	0	0	0	0
	Scenario 3	0	1	0	0	0	0	0	0	0	0	0

Appendix Table 3: Constraint set and candidate set of test for influence of asymmetric faithfulness constraints on learners

1. Asymmetric constraint set	2. Opposite asymmetric constraint set	3. Symmetric constraint set	Candidate set
*/õ/|o|	-	*/õ/|o|	[on] /on/ |on|
*/on/	*/o/|o|	*/o/|o|	[on] /õn/ |on|
*/õn/	*/on/	*/on/	[õn] /on/ |on|
*[on]/on/	*/õn/	*/õn/	[õn] /õn/ |on|
*[on]/õn/	*[on]/on/	*[on]/on/	
*[õn]/on/	*[on]/õn/	*[on]/õn/	
*[õn]/õn/	*[õn]/on/	*[õn]/on/	
	*[õn]/õn/	*[õn]/õn/	

Appendix Table 4: Results of test for influence of asymmetric faithfulness constraints on learners

		perception consistent learners |on|	consistent learners [õn]	production consistent learners |on|	perception [on] [on]/on/|on|	perception [on] [on]/õn/|on|	perception [õn] [õn]/on/|on|	perception [õn] [õn]/õn/|on|	production |on| [on]/on/|on|	production |on| [on]/õn/|on|	production |on| [õn]/on/|on|	production |on| [õn]/õn/|on|
asymmetric	dist1	1	0	0	27	73	31	69	0	0	31	69
	dist2	0	1	0	30	70	25	75	32	67	1	0
	dist3	0	1	1	31	69	29	71	17	37	14	32
xasymmetric	dist1	0	0	0	70	30	72	28	1	0	63	36
	dist2	0	0	0	74	26	66	34	57	42	0	1
	dist3	1	0	0	69	31	68	32	24	17	44	15
symmetric	dist1	0	0	0	61	39	48	52	0	1	36	63
	dist2	0	0	0	53	47	58	42	48	52	0	0
	dist3	0	0	0	51	49	53	47	22	31	25	22

Appendix Table 5: Results of test for influence of asymmetric faithfulness constraints on learners. In percentages per investigated group

constraints	perception [on] [on]/on/|on|	perception [on] [on]/õn/|on|	perception [õn] [õn]/on/|on|	perception [õn] [õn]/õn/|on|	production |on| [on]/on/|on|	production |on| [on]/õn/|on|	production |on| [õn]/on/|on|	production |on| [õn]/õn/|on|	/ON/PERC	/ÕN/PERC	/ON/PROD	/ÕN/PROD	/ON/TOT	/ÕN/TOT
asymmetric */õ/|o|	29%	71%	71%	29%	16%	35%	15%	34%	29%	71%	32%	68%	30%	70%
xasymmetric /o/|o|	71%	29%	30%	70%	27%	20%	36%	17%	70%	30%	63%	37%	68%	32%
symmetric */õ/|o| & /o/|o|	55%	45%	47%	53%	23%	28%	20%	28%	54%	46%	44%	56%	51%	49%
distribution1	53%	47%	53%	47%	0%	56%	43%	0%	52%	49%	44%	56%	49%	51%
distribution2	52%	48%	52%	48%	46%	54%	0%	0%	51%	51%	46%	54%	49%	51%
distribution3	50%	50%	50%	50%	21%	28%	28%	23%	50%	50%	49%	51%	50%	50%

Appendix Table 6: Number of input steps that lead to re-ranking (out of 100 total steps taken from second symmetry learning test in Section 4.1)

Learner with constraint set A	Learner with constraint set B
22	31

Appendix Table 7: Constraint re-ranking per constraints for one learner with constraint set A and one learner with constraint set B (out of 100 total steps taken from second symmetry learning test in Section 4.1)

	Learner with constraint set A		Learner with constraint set B	
	Ranked up	Ranked down	Ranked up	Ranked down
*/õ/|o|	3	4	4	4
*/o/|o|	-	-	4	4
*/on/	3	1	1	2
*/õn/	1	1	2	1
*/ot/	1	1	3	2
*/õt/	2	2	2	3
*[ot]/ot/	0	0	1	3
*[ot]/õt/	1	1	4	7
*[õt]/ot/	3	3	5	2
*[õt]/õt/	4	4	5	3
*[on]/on/	8	8	5	4
*[on]/õn/	5	5	5	2
*[õn]/on/	0	0	4	6
*[õn]/õn/	1	1	2	4
Total	32	31	47	47