# Can an AI learn Irish?

## The emergence of phonological features in Irish initial mutations

Ray Pelupessy
12608696

BA thesis Linguistics
University of Amsterdam
Supervised by:
Paul Boersma
27/6/2022

UNIVERSITY OF AMSTERDAM

**Abstract:**

This paper aims to show how a Neural Network processes and categorizes Irish initial mutations. After training, the network is able to produce and comprehend forms through the input layer, but not all forms are acquired correctly. The network has difficulty in both the comprehension and production of mutated forms. Any phonemes that can mutate are acquired less accurately than those that do not mutate. Though it is not able to correctly produce or comprehend, further analysis (in the form of a cosine similarity matrix) shows that the patterns formed in the network contain parallels to the mutation patterns. Regardless of these patterns, the network did not learn Irish initial mutations as intended. This would most likely require a more convoluted network design.

**Acknowledgements:**

I would like to thank Isabel for her continued support through this whole process, from helping me choose my topic to motivating my writing. She is my biggest inspiration and never fails to spark interest in the latest linguistic topic. I will never be able to express the linguistic inspiration she gives me in full, but I hope this thesis can serve as a glimpse into that inspiration.

# Table of Contents

# 1 Introduction

Initial Mutations (IMs) are changes at the start of a word that can be used to indicate changes in case, other morphosyntactic contexts, or simply occur due to phonological processes. In Irish, IMs are often described as being morphophonological. This is because they show some phonological regularity while still depending upon morphosyntactic cues to occur. Green (2006) argues that even though there was a historical precedent for the phonological classification, in Modern Irish this is no longer accurate and it would be more precise to describe Irish IMs as being purely morphological. To support this claim, Green uses an Optimality Theory analysis, but this is naturally not the only way to test this hypothesis. Therefore this paper will use a computational approach to this theory.

This computational approach is in fact based on a paper trying to disentangle phonetic qualities from morphological qualities (Boersma, Chládková & Benders, 2021). The authors use a deep Restricted Boltzmann Neural Network to analyze the morphophonological features of a toy-language, testing whether the NN could show a separation between patterns in the phonetics and patterns in the morphology. Because the model was bidirectional, when only sound was given as input it could predict the correct morphemes and semantic characteristics, and when only morphemes and semantic nodes were activated it would predict the correct sound output.

Combining the theory of Green (2006) with the NN of Boersma, Chládková & Benders (2021), the hypothesis can be made that a NN can determine whether Irish IMs are purely phonological, purely morphological, or a combination of both. To test this hypothesis, a model has to be created that can produce and comprehend Irish IMs in a manner similar to the toy language. This model should then be trained on data and analyzed to see what patterns it forms. The patterns should be representative of the nature of Irish IMs, furthermore showing whether the application of such a NN on real language data is comparable in performance to data from a toy language.

In 1.1 an outline of Irish initial mutations will be presented. In 1.2 the morphosyntactic aspects will be explained. Finally, in 1.3, an explanation of the neural network's design will be given.

## 1.1 Irish initial mutations

The Irish consonant inventory is shown in Table 1. Although much of the literature on Irish writes palatalized consonants as (for example) [p'] instead of [$p^j$], this could cause confusion for those unfamiliar with the literature. Therefore, in this paper, palatalization will simply be indicated with the /$^j$/ as is typical in IPA

*Table 1: The Irish consonant inventory, adapted from Hickey (2014)*

| | Labial | Alveolar | Velar | Glottal |
|---|---|---|---|---|
| Voiceless stops | p pʲ | t tʲ | k kʲ | |
| Voiced stops | b bʲ | d dʲ | g gʲ | |
| Voiceless fricatives | f fʲ | s sʲ | x xʲ | h |
| Voiced fricatives | v vʲ | | ɣ ɣʲ | |
| Nasals | m mʲ | n nʲ | ŋ ŋʲ | |
| Liquids | | l lʲ r rʲ | | |
| Semivowels* | w | | j | |

* = [w] is an allophone of [v] while [j] is an allophone of [ɣʲ]

In various literature discussing the initial mutations of Irish, it is usually described as a phonological change that later became morphosyntactic in nature (Hickey, 2014). This is because in Old Irish, IMs were much more regular compared to the IMs of Modern Irish. Though there is some variation in transcriptions of these IMs depending on the author, the IMs as used in this paper will generally follow the transcriptions of Green (2006) and Hickey (2014) as Hickey's documentation of Irish phonology was quite extensive, but the transcriptions by Green formed the basis for his hypothesis and should be applied to allow comparison of the results to his conclusions.

The mutations as described in Table 2 show the phonetic changes that occur with IMs based on the initial phoneme. As can be seen not every phoneme in the Irish inventory changes, most notably the /l/, /r/ and /n/. All consonants between slashes are underspecified for palatalization, while those between square brackets have the unpalatalized form on the left and the palatalized form on the right. Square brackets with two phonemes in them, i.e. [ɣ, j], show variation of these two consonants depending on the dialect and/or the following vowel. The exception to this is /s/ becoming /t/ under certain morphosyntactic conditions.

*Table 2: Irish initial mutations, adapted from Green (2006)*

| Initial Phoneme | Lenited phoneme | Eclipsed phoneme |
|---|---|---|
| /p/ | /f/ | /b/ |
| /t/ | [h] [h, $x^j$] | /d/ |
| /k/ | /x/ | /g/ |
| /b/ | [w, v] [$v^j$] | /m/ |
| /d/ | [ɣ] [$ɣ^j$, j] | /n/ |
| /g/ | [ɣ] [$ɣ^j$, j] | /ŋ/ |
| /f/ | Ø | [w, v] [$v^j$] |
| /s/ | [h, t] [h, $t^j$, $x^j$] | no change |
| /m/ | [w, v] [$v^j$] | no change |

As can be seen, Irish has two separate processes, one called "Lenition" and another called "Eclipsis" (thought these names may differ per author). Eclipsis is considered the more phonologically consistent of the two, as it turns voiceless plosives into voiced plosives and voiced plosives into nasals. The only exception to this is /f/, which is a fricative rather than a plosive, but it still follows the [voiceless] → [voiced] process.

Lenition is much more phonologically irregular, but is generally described as turning consonants into fricatives, while deleting fricatives in most cases. Lenition shows greater variety in the output forms, as the output for /t/, /b/, /m/, /d/, /g/, and /s/ are all dependent on the palatalization of the consonant. There is also a merger of /b/ and /m/ to /v/, and a merger of /d/ and /g/ to /ɣ/.

One aspect that could impact the analysis of the data by a NN, is that not all consonants in the Irish phonetic inventory are allowed in the onset in all situations. Table 3 lists the phonemes as *independent* if they occur in the onset when no IM is applied, or *dependent* if they can only occur in the onset after an IM is applied (Hickey, 2014).

*Table 3: Independent and dependent consonants (underspecified for palatalization)*

| Fully Independent | Partially Independent | Dependent |
|---|---|---|
| /p/ /s/ /k/ /l/ /r/ | /f/ /b/ /m/ /t/ /d/ /n/ /g/ | /v/ /ɣ/ /x/ /h/ |

The "Fully Independent" vs. "Partially Independent" division was not made in the source (Hickey, 2014) but could still prove an important distinction, as the fully independent consonants only occur in onsets where no mutation has occurred, while partially independent consonants occur both in regular onsets and in onsets where a mutation has taken place. The fully independent phonemes thus appear in more phonologically consistent environments than the partially independent phonemes do.

## 1.2 Irish morphosyntax

The morphosyntactic qualities of IMs are important to mention, as these will form a significant part of the NN. The first quality covered here is grammatical gender of the noun. This gender is inherent to the noun and will therefore not be part of the input in the NN. Noun gender does, however, have an influence on which mutations the noun takes on in different contexts, and therefore will still be noted in the data.

Other factors are case, definiteness, and plurality. Plurality and definiteness are the most straightforward, as Irish makes a binary distinction between singular and plural, and between definite and indefinite. The plurality distinction can usually be noted through a change in the ending of the noun. Varying suffixes are used to form plurals, but this will not be elaborated upon as this is not relevant to the hypothesis presented here. Definiteness is occasionally reflected on the noun in the form of initial mutations, but is more clearly visible in the addition of “an” or “na” before the noun. Indefinite nouns lack any sort of determiner.

In Irish there are generally considered to be four cases: nominative, genitive, prepositional/dative, and the vocative. Though the accusative does exist, it is only noticeable in the pronouns, which are not relevant for this paper as they are rarely subject to initial mutation. Other than pronouns, the accusative takes the nominative forms. The indefinite nominative is, as in many languages, considered the “base” form, while processes such as IM and palatalization are applied to create the other cases. The genitive occurs with possessives and a few prepositions, and is often formed by applying lenition or eclipsis and palatalizing the final consonant of the word. The dative occurs only when a preposition is used, and usually causes lenition of the noun. Finally, the vocative is used when addressing someone directly, and is formed through lenition.

## 1.3 Design of the neural network

The first design of the NN as seen in Image 1 is based on the Boersma et. al (2021) paper, though some adaptations were made to make it more suited to the Irish data.

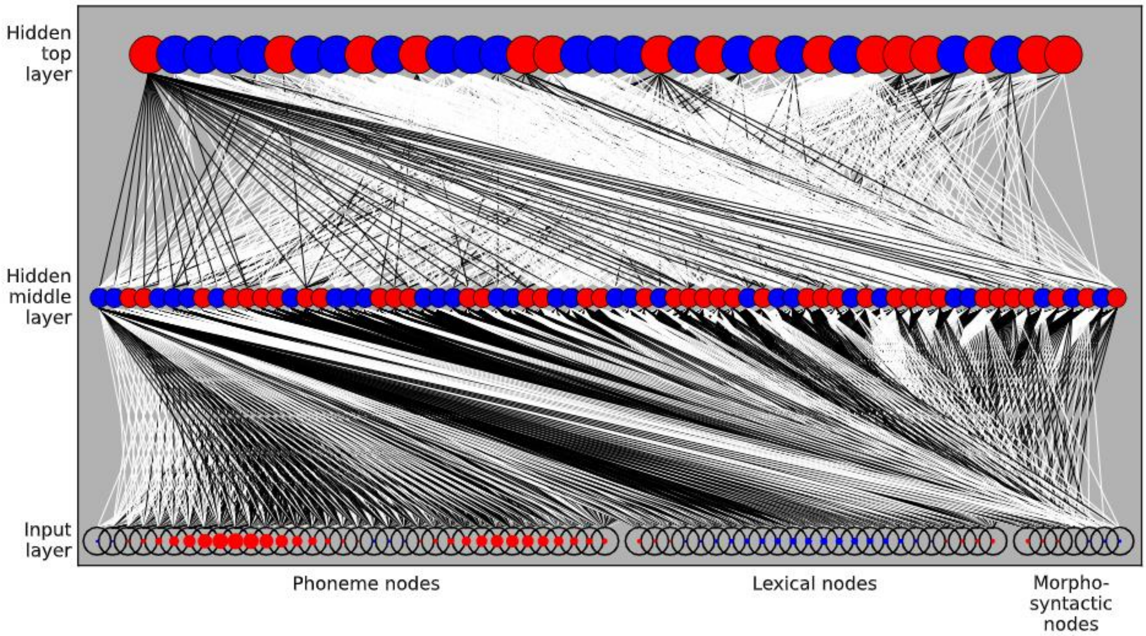


*Image 1: An initial design of the NN*

As can be seen in Image 1, there are three “groups” of nodes that are based on the information as described in sections 1.1 and 1.2. The leftmost group of nodes represent the initial phonemes, which is also the part that differs most strongly from (Boersma, Chládková & Benders, 2021), as they used a  spectral frequency-based input, rather than the individual phonemes. This was unfortunately not realistic for the consonants however, leading to this change in the design where each individual node represents a phoneme. There is also a node for a null-consonant (Ø), as this needed to be included for the /f/ lenition. If this node were not included, it might cause the network to not analyze this sound change in a similar manner to the others.

The middle group represents the meanings of the words with their orthographic representation. The form used is the nominative indefinite singular, as this always uses the unmutated initial consonant. While this may be more confusing than simply using the IPA transcription, the orthographic form is only representative of some of the forms, since palatalization of the final consonant is not included in order to limit the number of nodes. It is best to view the lexical nodes as being the underlying representation of the word and its meaning, rather than the actual form expected in every situation.

The rightmost group of nodes represents the morphosyntax. It includes the plural, singular, definite, and indefinite characteristics. Case is limited to nominative, genitive, and dative, as these cases together are already representative of all the sound changes. The vocative shows many similarities to the genitive, while the accusative is almost fully merged with the nominative case.

All of this together means that the first layer has 89 nodes: 34 for the phonemes, 48 for the individual words, and 7 for the morphosyntactic nodes. The second layer, which processes the input first, is made up of 70 nodes, while the third layer consists of 35 nodes. The number of nodes in the second and third layers were arbitrarily chosen; depending on the performance of this NN it might be possible to see what difference the number of nodes in each layer makes on the outcomes.

Due to the large number of input nodes, the actual visualization in Image 2 does not include node labels but rather numbers (0 through 88). This is to reduce visual clutter of the visualization. Appendix A contains the numbers each phoneme, lexical entry, and morphosyntactic feature correspond to.

# 2. Methodology

The code written for the NN can be found in the additional files on the Archive of the Institute of Phonetic Sciences (IFA) Amsterdam. The code was based upon various works (Boersma, Chládková & Benders, 2021; Terlou, 2021; Meijer, 2021), with the Python code provided by Terlou serving as the basis of the full model. Any changes between the original code as seen in (Terlou, 2021) and the code as applied in this thesis are the author's.

The input data for the NN can be found in Appendix B. Each entry contains a word with each initial phoneme being represented in the data. Two words with an independent initial phoneme were selected and the forms in the nominative, genitive, and dative were all recorded in both singular and plural, and definite and indefinite forms (Foras na Gaeilge, 2022). Training was performed using the complete dataset.

The training of the NN occurs in several steps. First the initialization of the model forms all the nodes and their connections, setting all weights to 0. Then, a data entry is chosen randomly and the corresponding nodes are activated, allowing the actual learning to occur. All formulas are taken from Boersma, Chládková & Benders (2021).

## 2.1 Settling phase

Each learning step starts off with the settling phase. This phase keeps the input layer "clamped," meaning that the input layer is static: active nodes stay active and inactive nodes remain inactive. The nodes on the input layer spread their activations, resulting in a change in the activations on all levels except the input layer. This happens through the functions as seen in (1) and (2), where the sigma represents the sigmoid function seen in (3).

(1) $y_l \leftarrow \sigma(b_l + \sum_{k=1}^{K} x_k u_{kl} + \sum_{m=1}^{M} z_m v_{lm})$

(2) $z_m \leftarrow \sigma(c_m + \sum_{l=1}^{L} y_l v_{lm})$

(3) $\sigma(x) := \frac{1}{1 + \exp(-x)}$

The functions use $x$, $y$, and $z$ to represent the nodes of the input layer, middle layer, and top layer respectively. $K$, $M$, and $L$ refer to the number of nodes, with $x_k$, $y_l$, and $z_m$ referring to the activity of any one of the specific nodes in that layer. The $b$ and $c$ reflect the bias of the current node, while the $u$ and $v$ reflect the weights between the two nodes $l$ and $k/m$.

In conclusion, (1) and (2) show that the excitation for each node is calculated by adding the existing bias to the sum of excitations from all active nodes connected to it. Nodes with a stronger connection will have a stronger influence on each other, but all excitations must remain within the range of 0 to 1 due to the sigmoid function in (3).

The steps of (1) and (2) are repeated 10 times before the network is assumed to have a somewhat stable state. This way the second phase can be initiated.

## 2.2 Hebbian learning phase

Following the settling phase is Hebbian learning. This phase involves changing the biases of nodes positively if they are activated, and increasing the connection weight of a connection if both nodes are activated simultaneously. The following functions (4-8) represent the manner in which this happens, with (4), (5), and (6) describing changes in bias for each of the layers, while (7) and (8) show the changes in connections between nodes.

(4) $a_k \leftarrow a_k + \eta x_k$

(5) $b_l \leftarrow b_l + \eta y_l$

(6) $c_m \leftarrow c_m + \eta z_m$

(7) $u_{kl} \leftarrow u_{kl} + \eta x_k y_l$

(8) $v_{lm} \leftarrow v_{lm} + \eta y_l z_m$

For each function, the current bias/weight is replaced by the addition of the bias/weight to any activation that occurs multiplied with the learning rate $\eta$ ($=0.001$).

## 2.3 Dreaming phase

The dreaming phase allows the network to interpret the input by "unclamping" the input layer, letting the input nodes change their activation states. To do this, activation should now spread to the lowest layer from the middle, which is done similarly as in (1) and (2). The formula can be seen in (9). No sigmoid function is used here.

(9) $x_k \leftarrow a_k + \sum_{l=1}^{L} u_{kl} y_l$

Following the first spread of activation, new activations are calculated for the second and top layer as seen in (10) and (11).

(10) $z_m \sim \mathcal{B}(\sigma(c_m + \sum_{l=1}^{L} y_l v_{lm}))$

(11) $y_l \sim \mathcal{B}(\sigma(b_l + \sum_{k=1}^{K} x_k u_{kl} + \sum_{m=1}^{M} z_m v_{lm}))$

These functions differ somewhat from the settling phase, as the Bernoulli distribution is used to introduce a degree of randomness to the learning. Taking the output of the sigmoid function as *p*, $z \sim \mathcal{B}(p)$ will either result in a 1 with a probability *p*, or 0 with a probability of 1 – *p*. The functions (9) through (11) are performed 10 times before the activities are assumed to be stable.

### 2.4 Anti-Hebbian learning

Finally, the model rounds out a single learning step with anti-Hebbian learning. The functions in (12) to (16) are the same as (4) to (8), but uses subtraction instead of addition to unlearn some of the patterns it has formed.

(12) $a_k \leftarrow a_k - \eta x_k$

(13) $b_l \leftarrow b_l - \eta y_l$

(14) $c_m \leftarrow c_m - \eta z_m$

(15) $u_{kl} \leftarrow u_{kl} - \eta x_k y_l$

(16) $v_{lm} \leftarrow v_{lm} - \eta y_l z_m$

In total, these steps conclude a single learning step, which will be performed for each data entry. To determine the patterns in phonology and morphology/morphosyntax, separate models will be trained for the phonology, morphosyntax, and a combined model that includes both will be trained last to determine any overlap in the morphophonology.

## 3. Training the model

After the Neural Network was created, it had to be trained and tested in various ways to determine its success in reproducing input and predicting forms that were left out of training. This training was done by letting the network perform 100.000 training steps on the training data, and then testing it on the relevant output. Each training step involved training on a randomly chosen entry in the dataset, while testing occurred only on specific selected forms (with the exception of 3.1). Of course, the results can vary somewhat due to the random selection and ordering of training data, so all results are from 100 "virtual learners" that were separately trained and tested. All this data was then saved and averaged over these 100 learners.

In 3.1 the full dataset was used in training to see if the model would reproduce these forms faithfully. From 3.2 onward, parts of the data are left out to see if it can correctly predict unseen forms. In 3.2, the data for half of the words is left out, with the exception of a single case, testing the network on its ability to extrapolate the nominative, genitive, and dative forms. In 3.3 the same is done but with the singular and plural. In 3.4 this is done with the definite and indefinite forms.

## 3.1 Initial testing

After training the model on the complete dataset, it was tested on both comprehension and production. The comprehension tests were done by activating a lexical node and a corresponding phoneme. This activation is then spread like in the settling phase of training, after which related morphosyntactic nodes should activate. Production tests were done by activating a lexical node alongside the morphosyntactic nodes, which results in the activation of a phoneme node.

### 3.1.1 Production

To test if the network is correctly able to produce phonemes, it was first fully trained on all forms. The network, seen in Image 2, went through 100.000 training steps, meaning that each datapoint was (on average) repeated 174 times, though this varied per learner, as the datapoint was determined randomly every training step. In Image 2, white lines indicate a negative connection, while black lines indicate a positive connection between nodes. None of the nodes are activated, as can be seen in the input layer by the lack of coloring.

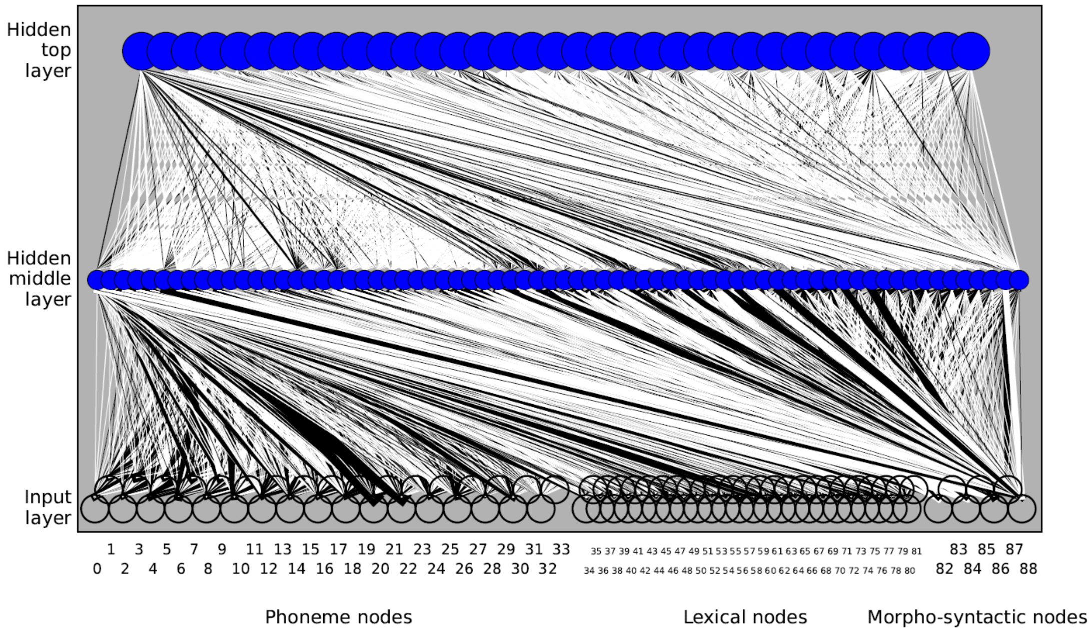


*Image 2: The Network after 100.000 training steps without any activated nodes.*

The model was tested on each of the datapoints it had been trained on by only providing the lexical entry and its morphosyntactic features. The activation was spread through the model 10 times while keeping the lexical and morphosyntactic nodes clamped so no changes could occur in the input. Most of the time multiple nodes would be activated in the output, therefore only the node with the highest activation level was selected as the actual output.

Table 4 shows the results of the Neural Network's performance, averaged over 100 learners. Notably, there are only ten phonemes that are produced correctly less than 50% of the time. This stands in contrast to twelve phonemes that are produced correctly more than 90% of the time.

*Table 4: Percentage of correct answers for the production of each phoneme and its occurrence in the dataset*

| Phoneme | Correctness | Occurrences (of 576) | Phoneme | Correctness | Occurrences (of 576) |
|---|---|---|---|---|---|
| Ø | 23.1% | 2.1% | n | 91.0% | 4.5% |
| p | 99.9% | 2.4% | nʲ | 92.0% | 4.5% |
| pʲ | 99.1% | 2.4% | l | 98.3% | 4.2% |
| b | 75.6% | 3.1% | lʲ | 98.4% | 4.2% |
| bʲ | 73.9% | 3.1% | r | 98.8% | 4.2% |
| f | 57.3% | 3.5% | rʲ | 99.5% | 4.2% |
| fʲ | 58.4% | 3.5% | k | 99.3% | 2.4% |
| v | 4.7% | 2.8% | kʲ | 98.9% | 2.4% |
| vʲ | 5.1% | 2.8% | g | 66.8% | 3.1% |
| m | 77.6% | 3.8% | gʲ | 71.1% | 3.1% |
| mʲ | 79.2% | 3.8% | x | 0.0% | 1.0% |
| t | 78.2% | 3.8% | xʲ | 0.3% | 1.0% |
| tʲ | 89.6% | 3.5% | ɣ | 19.6% | 1.7% |
| d | 88.9% | 3.5% | ɣʲ | 17.2% | 1.7% |
| dʲ | 88.0% | 3.5% | ŋ | 0.0% | 0.7% |
| s | 99.0% | 2.8% | ŋʲ | 0.0% | 0.7% |
| sʲ | 99.5% | 3.1% | h | 1.8% | 2.8% |

As can be seen in Table 4, the phonemes /v/, /x/, /ŋ/, and /h/ were produced incorrectly in (nearly) all cases. In 1.1 it was mentioned that these are the dependent phonemes, which can only occur in a mutated environment. While for /x/ and /ŋ/ it could be argued that their lack of frequency in the dataset might be a reason for their low correctness percentage, this argument is quickly debunked by the frequency of /h/ in the dataset (2.8%), which is higher than other phonemes that are produced correctly much more frequently. While the /Ø/ does not officially fall into any of the dependency categories, in the dataset it only occurs as the mutated form of /f/. Since the /Ø/ is less prevalent in the dataset, yet produced correctly far more than /h/, this further debunks the frequency argument.

Additionally, the [ɣ] and [ɣʲ] each occurred in only 1.7% of the dataset, yet were produced correctly more frequently than the other dependent phonemes. This stark contrast can be partially explained due to the occurrence of /ɣ/ in both the lenition of /d/ and /g/, as this would most likely produce more connections, leading to more activation. However, the

/v/ occurs in the lenition of both /b/ and /m/, yet showed less correct production when compared to /ɣ/, so another factor could be at play in the success of the production of /ɣ/.

On the opposite end of the spectrum are the fully independent phonemes, which can only occur in non-mutated positions. The fully independent phonemes (/p/ /s/ /k/ /l/ /r/) were all produced correctly over 98.3% of the time, showing that the network had no trouble acquiring these phonemes.

Finally, the partially independent phonemes were produced correctly in most situations but not as overwhelmingly as the fully independent phonemes. Further data analysis may show that the cases of incorrect production were those where the phoneme was mutated, while the correct productions were those where the phoneme was unmutated. This would fall in line with the Network's tendency to incorrectly produce dependent phonemes.

### 3.1.2 Comprehension

For testing comprehension, the model was presented only with a phoneme and a related lexical node. The output was then decided per category: the highest activated case node, plurality node, and definiteness node. Overall the model performed much more poorly on comprehension than production, as can be seen from the results in Table 5.

*Table 5: Percentage of correct answers for comprehension of morphosyntax*

| **Morphosyntactic feature** | | **Correctness (stdev)** |
|---|---|---|
| **Case** | Nominative | 21.0% (56.4) |
| | Genitive | 32.2% (32.2) |
| | Dative | 60.6% (67.7) |
| **Plurality** | Singular | 54.2% (84.6) |
| | Plural | 47.3% (84.7) |
| **Definiteness** | Indefinite | 64.6% (81.2) |
| | Definite | 40.2% (83.2) |

However, this analysis was not done exactly like the production analysis. While for production the desired phoneme was directly compared to the actual output, for comprehension this was not entirely fair. Since phonemes can occur in various morphosyntactic environments, a different analysis had to be performed. To illustrate, take the phoneme /b/ and the lexical node 'cow'. This entry could be associated (correctly) with the nominative singular indefinite form, the genitive singular (in)definite form, the nominative plural (in)definite form, the genitive plural indefinite form, and finally the dative plural definite form. Because there are many possible options for this single phoneme-lexeme combination, neither the network nor a real person would be able to decide the correct case-plurality-definiteness combination based solely on this information.

For this reason, it was better to compare the percentage of occurrences of each phoneme-case pairing (per lexical item) to see if the averages matched up, rather than the harsher criteria of needing each testing item to have the correct output. So if the Network guesses 15/70 occurrences of /b/ plus 'cow' to be the nominative, but the input data contains the nominative for 35 of 70 occurrences, those 15 occurrences are still counted as correct even if they were supposed to be assigned a different case. This way the Network is not penalized for mistakes that are, by definition, impossible to predict.

In Table 5, the results of such an analysis are shown. The dative, singular, and indefinite forms are produced most reliably, while the nominative case is produced correctly just over a fifth of the time. Although the results aren't as promising as the results of production, this is logical, as there is less information to process and draw conclusions from. The data showed a tendency to produce the same morphosyntactic features for any phoneme-lexeme pair, which makes sense, as humans would most likely also assume the most common form is being used when presented with a word-form. If someone is shown the word "bó" ('cow') they are more likely to assume it as the nominative singular indefinite form, rather than the genitive singular or plural indefinite. In this sense, the network may be showing more realistic behavior than was tested for.

## 3.2 Case predictions

After running the tests in 3.1, the neural network had to be tested on unseen data to test if it is able to extrapolate patterns to new data. To this end, for every item in the training data, the forms of half the dataset were reduced by one case. This way the two lexical entries for each phoneme are still trained, but one of them is lacking the entries for one of the three cases, which will instead be used as the testing data. Since this resulted in a smaller training dataset, each datapoint was iterated over 208 times on average, though there was still random selection of datapoints during training just like in 3.1. Testing was done only on the unseen data, corresponding to the odd numbered entries in the table in Appendix A. Full paradigms for these words can be found in Appendix B.

### 3.2.1 Nominative

3.2.1.1 Production

For the nominative case, not all possible phonemes can be found in production. Those that were present in the data, however, were produced according to the data in Table 6.

*Table 6: Percentage of correct answers for the production of each phoneme in the nominative case*

| Phoneme | Correctness | Occurrences (of 96) | Phoneme | Correctness | Occurrences (of 96) |
|---|---|---|---|---|---|
| **Ø** | 20.0% | 1.0% | **dʲ** | 85.8% | 4.2% |
| **p** | 55.0% | 4.2% | **n** | 100% | 4.2% |
| **pʲ** | 52.8% | 4.2% | **nʲ** | 98.5% | 4.2% |
| **b** | 96.0% | 3.1% | **l** | 100% | 4.2% |
| **bʲ** | 94.7% | 3.1% | **lʲ** | 99.8% | 4.2% |
| **f** | 68.3% | 3.1% | **r** | 99.0% | 4.2% |
| **fʲ** | 5.0% | 4.2% | **rʲ** | 99.5% | 4.2% |
| **v** | 5.0% | 1.0% | **k** | 66.8% | 4.2% |
| **vʲ** | 2.0% | 1.0% | **kʲ** | 60.5% | 4.2% |
| **m** | 71.5% | 4.2% | **g** | 12.5% | 4.2% |
| **mʲ** | 73.8% | 4.2% | **gʲ** | 8.0% | 4.2% |
| **t** | 78.0% | 5.2% | **s** | 99.0% | 3.1% |
| **tʲ** | 92.5% | 4.2% | **sʲ** | 97.8% | 4.2% |
| **d** | 86.8% | 4.2% | | | |

Once again, the dependent and partially independent phonemes are produced correctly fewer times than the independent phonemes are. Most notably, /v/ and /g/ are produced correctly least of all phonemes, along with [fʲ], which seems to correspond with the data from section 3.1.

However, it is interesting to note that, other than /Ø/, /v/, and /t/, none of the phonemes appearing in the nominative case are mutated in the nominative. The mutations of the aforementioned phonemes only occurs in the definite forms. Thus all partially independent phonemes here are unmutated forms with the exception of /t/.

3.2.1.2 Comprehension

The results of comprehension can be seen in Table 7. Immediately it is interesting to see how the nominative is comprehended correctly only 1.5% of the time. This can be explained in several ways, but it is logical to assume that because this data was not in the training data, leading the model to struggled with accurately associating nominative data with the correct case.

*Table 7: Percentages of correct answers for comprehension of morphosyntax in the nominative case*

| Morphosyntactic feature | | Correctness (stdev) |
|---|---|---|
| **Case** | Nominative | 1.5% (11.9) |
| **Plurality** | Singular | 47.0% (34.6) |
| | Plural | 54.0% (34.5) |
| **Definiteness** | Indefinite | 52.6% (34.6) |
| | Definite | 46.7% (34.6) |

Another notable difference between this data and the data from Table 5, is that the singular is produced less faithfully than the plural. This was not the case for the full dataset, which means that either the plural is more regular in the nominative case, or the model performed similarly and this difference is simply coincidental. Looking at the testing data, we find that the plural is indeed more regular than the singular, with no mutations in the nominative plural at all. Although there are only a few mutations in the nominative singular (4 out of 96), this small amount is enough to cause a difference in the comprehension. The indefinite forms are still produced correctly more often than definite forms, which makes sense when we see that the test data included no mutations in the indefinite, but some in the definite datapoints (4 out of 96).

### 3.2.2 Genitive

#### 3.2.2.1 Production

Table 8 shows the production of phonemes in the genitive, which includes more of the dependent phonemes than the nominative, but still lacks the /h/.

*Table 8: Percentage of correct answers for the production of each phoneme in the genitive case*

| Phoneme | Correctness | Occurrences (of 96) | Phoneme | Correctness | Occurrences (of 96) |
|---|---|---|---|---|---|
| Ø | 5.0% | 1.0% | n | 80.0% | 5.2% |
| p | 95.5% | 2.1% | n$^j$ | 80.0% | 5.2% |
| p$^j$ | 100% | 2.1% | l | 99.3% | 4.2% |
| b | 49.3% | 4.2% | l$^j$ | 99.5% | 4.2% |
| b$^j$ | 51.0% | 4.2% | r | 98.3% | 4.2% |
| f | 27.3% | 4.2% | r$^j$ | 98.3% | 4.2% |
| f$^j$ | 57.7% | 3.1% | k | 100% | 2.1% |
| v | 2.0% | 2.1% | k$^j$ | 100% | 2.1% |
| v$^j$ | 1.5% | 2.1% | g | 62.0% | 3.1% |
| m | 71.3% | 4.2% | g$^j$ | 65.7% | 3.1% |
| m$^j$ | 72.8% | 4.2% | x | 0.0% | 1.0% |
| t | 97.3% | 3.1% | x$^j$ | 0.0% | 1.0% |
| t$^j$ | 74.0% | 4.2% | ɣ | 3.0% | 1.0% |
| d | 70.5% | 4.2% | ɣ$^j$ | 1.0% | 1.0% |
| d$^j$ | 67.75% | 4.2% | ŋ | 0.0% | 1.0% |
| s | 89.3% | 4.2% | ŋ$^j$ | 0.0% | 1.0% |
| s$^j$ | 98.0% | 3.1% | | | |

It is clear /Ø/, /v/, /x/, /ɣ/, and /ŋ/ are produced incorrectly far more often than other phonemes. While [f] is produced incorrectly more than [f$^j$], [f] actually occurs in a non-mutated environment more than [f$^j$] does. A similar discrepancy can be seen between [t] and [t$^j$], where [t] only occurs in non-mutated environments, but [t$^j$] is the mutated counterpart of [s$^j$], which is reflected in the correctness. All phonemes with a higher correctness percentage are those which show few mutations in the genitive, indicating that the network was unable to extrapolate the patterns for the mutations here.

3.2.2.2 Comprehension

Table 9 shows the results for comprehension on the genitive testing data. Just like in the nominative, the genitive itself is produced correctly very rarely, supporting the hypothesis from 3.2.1.2 that the network struggles to produce the case it is being tested on. This is most likely because the data wasn't present during training, thus leading the connections to be weaker.

*Table 9: Percentages of correct answers for comprehension of morphosyntax in the genitive case*

| **Morphosyntactic feature** | | **Correctness (stdev)** |
|---|---|---|
| **Case** | Genitive | 2.2% (14.4) |
| **Plurality** | Singular | 54.2% (34.5) |
| | Plural | 46.3% (34.5) |
| **Definiteness** | Indefinite | 59.5% (34.0) |
| | Definite | 37.8% (33.6) |

Once again comparing the results to the testing data, there is a tendency to comprehend the singular and indefinite more faithfully. The singular has fewer mutations than the plural (10 vs 14 respectively), while the indefinite has no mutations at all. The definite forms have both lenited and eclipsed forms (24 total) and together with the data from the nominative the hypothesis can be proposed that unmutated forms are comprehended more accurately.

### 3.2.3 Dative

#### 3.2.3.1 Production

As with the other cases, the production results are represented in Table 10. The dative contains all possible initial phonemes, unlike the nominative and genitive.

*Table 10: Percentage of correct answers for the production of each phoneme in the dative case*

| Phoneme | Correctness | Occurrences (of 96) | Phoneme | Correctness | Occurrences (of 96) |
|---|---|---|---|---|---|
| Ø | 0.5% | 4.2% | n | 99.0% | 4.2% |
| p | 99.0% | 1.0% | $n^j$ | 96.8% | 4.2% |
| $p^j$ | 100% | 1.0% | l | 97.8% | 4.2% |
| b | 49.0% | 2.1% | $l^j$ | 94.5% | 4.2% |
| $b^j$ | 50.0% | 2.1% | r | 99.5% | 4.2% |
| f | 33.0% | 3.1% | $r^j$ | 99.0% | 4.2% |
| $f^j$ | 32.7% | 3.1% | k | 100% | 1.0% |
| v | 0.0% | 5.2% | $k^j$ | 100% | 1.0% |
| $v^j$ | 0.6% | 5.2% | g | 50.0% | 2.1% |
| m | 66.0% | 3.1% | $g^j$ | 50.0% | 2.1% |
| $m^j$ | 65.7% | 3.1% | x | 0.0% | 2.1% |
| t | 66.0% | 3.1% | $x^j$ | 0.0% | 2.1% |
| $t^j$ | 97.0% | 2.1% | ɣ | 0.0% | 4.2% |
| d | 99.0% | 2.1% | $ɣ^j$ | 0.0% | 4.2% |
| $d^j$ | 99.0% | 2.1% | ŋ | 0.0% | 1.0% |
| s | 99.0% | 1.0% | $ŋ^j$ | 0.0% | 1.0% |
| $s^j$ | 99.0% | 2.1% | h | 0.0% | 8.3% |

What stands out in the above data, is the consistent failure to produce the dependent phonemes once again. Although the dependent phonemes are equally present in the data as they were in other cases, the correctness has dropped to 0.0% for nearly all, with [$v^j$] and /Ø/ being 0.6% and 0.5% respectively. This is interesting, as Table 5 indicated that, if fully trained, the model is best able to identify the dative case from phonemic input, though this could mean that the model is simply very accurate in learning the patterns of independent phonemes in the dative. That hypothesis is supported by the way in which the percentages nearly perfectly match the patterns of mutation in these phonemes. While /p/, /t/, /d/, /s/, /n/, /l/, /r/, and /k/ only occur in unmutated conditions in the dative, /b/ and /g/ each occur in mutated conditions half the time, and unmutated conditions in the other half of instances. /f/ and /m/ occur in mutated conditions in one third or two third of instances respectively.

This division shows clearly that in this scenario, any patterns of mutation are not applied to the unseen data. The unmutated forms are produced near-perfectly, but any correct productions in mutated forms appear more coincidental than they do intentional.

3.2.3.2 Comprehension

The results of the comprehension test of the dative are shown in Table 11. Compared to the nominative and genitive, the dative is comprehended much more accurately. However, compared to the full testing, there is still a large discrepancy, once more supporting the hypothesis that whatever feature is being trained on will be comprehended much less accurately.

*Table 11: Percentages of correct comprehension of morphosyntax in the dative case*

| **Morphosyntactic feature** | | **Correctness (stdev)** |
|---|---|---|
| **Case** | Dative | 22.6% (41.0) |
| **Plurality** | Singular | 58.1% (34.2) |
| | Plural | 42.2% (34.2) |
| **Definiteness** | Indefinite | 54.6% (34.5) |
| | Definite | 54.6% (34.5) |

The results of this test are furthermore notable in the area of definiteness, where both indefinite and definite forms are produced correctly more than they are produced incorrectly. Even though the definite contains 11 mutated forms and the indefinite 36 lenited forms, the comprehension is not weakened by the prevalence of mutated phonemes. This debunks the hypothesis that mutated forms are comprehended less accurately, instead lending credence to the hypothesis that some degree of mutation can aid comprehension. This is further supported by the stronger presence of mutation in the singular as compared to the plural. Notable is that the singular also has both eclipsis and lenition, rather than only lenition in the plural, which may mean that the variation of initial phonemes may instead be what aids comprehension.

## 3.3 Plurality

For plurality, the data was once again split like with the cases, though since there are only two features rather than three with case, there is more testing data. 144 items were used for testing, while the remaining 432 items were used for training. Thus in 100,000 training steps, each item is repeated approximately 231 times. In 3.3.1 only the singular forms of the odd numbered entries in Appendix A were tested, while in 3.3.2 the plural of those same words was tested.

### 3.3.1 Singular

3.3.1.1 Production

Table 12 shows the production of all the singular forms. All possible initial phonemes appear in the singular, though not all were produced faithfully.

*Table 12: Percentage of correct answers for the production of each phoneme in the singular*

| Phoneme | Correctness | Occurrences (of 144) | Phoneme | Correctness | Occurrences (of 144) |
|---|---|---|---|---|---|
| Ø | 2.0% | 2.8% | n | 96.0% | 4.2% |
| p | 98.0% | 2.1% | nʲ | 94.3% | 4.2% |
| pʲ | 97.7% | 2.1% | l | 94.5% | 4.2% |
| b | 74.0% | 2.8% | lʲ | 97.0% | 4.2% |
| bʲ | 73.8% | 2.8% | r | 96.2% | 4.2% |
| f | 55.6% | 3.5% | rʲ | 94.7% | 4.2% |
| fʲ | 56.6% | 3.5% | k | 99.3% | 2.1% |
| v | 3.0% | 3.5% | kʲ | 99.7% | 2.1% |
| vʲ | 2.6% | 3.5% | g | 73.3% | 2.8% |
| m | 78.6% | 3.5% | gʲ | 74.0% | 2.8% |
| mʲ | 76.6% | 3.5% | x | 0.0% | 1.4% |
| t | 65.9% | 4.9% | xʲ | 0.0% | 1.4% |
| tʲ | 77.8% | 4.2% | ɣ | 10.3% | 2.1% |
| d | 89.6% | 3.5% | ɣʲ | 8.3% | 2.1% |
| dʲ | 92.0% | 3.5% | ŋ | 0.0% | 0.7% |
| s | 96.3% | 2.1% | ŋʲ | 0.0% | 0.7% |
| sʲ | 95.5% | 2.8% | h | 13.8% | 2.8% |

Notable is the percentages of correct answers for /ɣ/ as compared to the other dependent phonemes. When compared to the tests on case, /ɣ/ is produced more faithfully, while /v/, /x/, and /ŋ/ are still not produced nearly as accurately.

3.3.1.2 Comprehension

The results in Table 13 show more support for the hypothesis that the feature being tested on is produced more poorly. With 6.4% the performance is much lower than what was seen in the full training of 3.1.

*Table 13: Percentage of correct answers for comprehension of morphosyntax in the singular*

| Morphosyntactic feature | | Correctness (stdev) |
|---|---|---|
| **Case** | Nominative | 67.6% (28.6) |
| | Genitive | 24.8% (29.9) |
| | Dative | 67.6% (32.4) |
| **Plurality** | Singular | 6.4% (29.4) |
| **Definiteness** | Indefinite | 60.4% (41.5) |
| | Definite | 43.3% (42.0) |

Surprisingly, the nominative and dative are both comprehended much better than the genitive. This does not line up with the results from 3.1, where the nominative was only comprehended correctly in 21% of items, and does not follow the adjusted hypothesis that mutations cause easier comprehension. Only 4 of the nominative forms are lenited, while the dative has an even division of eclipsed, lenited, and unmutated forms. However, they are still performed equally well on. This indicates a more detailed analysis would be required to determine the cause of success in the correct comprehension of these forms.

The same can be said about the definite and indefinite forms, where indefinite forms only show lenition, but definite forms have both lenition and eclipsis. While indefinite forms are produced more accurately, this may be caused by reasons other than purely which mutations occur.

### 3.3.2 Plural

3.3.2.1 Production

Just like in the singular, all possible initial phonemes are present in the plural. They were produced according to the data in Table 14.

*Table 14: Percentage of correct answers for the production of each phoneme in the plural*

| Phoneme | Correctness | Occurrences (of 144) | Phoneme | Correctness | Occurrences (of 144) |
|---|---|---|---|---|---|
| Ø | 52.0% | 1.4% | n | 84.4% | 4.9% |
| p | 76.8% | 2.8% | nʲ | 84.1% | 4.9% |
| pʲ | 78.3% | 2.8% | l | 97.8% | 4.2% |
| b | 61.0% | 3.5% | lʲ | 98.3% | 4.2% |
| bʲ | 58.2% | 3.5% | r | 98.2% | 4.2% |
| f | 46.0% | 3.5% | rʲ | 96.5% | 4.2% |
| fʲ | 42.4% | 3.5% | k | 84.3% | 2.8% |
| v | 21.0% | 2.1% | kʲ | 84.5% | 2.8% |
| vʲ | 19.6% | 2.1% | g | 45.2% | 3.5% |
| m | 69.5% | 4.2% | gʲ | 41.6% | 3.5% |
| mʲ | 66.8% | 4.2% | x | 10.0% | 0.7% |
| t | 98.8% | 2.8% | xʲ | 16.0% | 0.7% |
| tʲ | 98.0% | 2.8% | ɣ | 40.5% | 1.4% |
| d | 76.6% | 3.5% | ɣʲ | 40.5% | 1.4% |
| dʲ | 79.8% | 3.5% | ŋ | 0.0% | 0.7% |
| s | 74.0% | 3.5% | ŋʲ | 0.0% | 0.7% |
| sʲ | 88.8% | 3.5% | h | 15.8% | 2.8% |

The contrast between this data and that of the singular is very noticeable when looking at the dependent phonemes. There is an improvement in the production of all phonemes except /ŋ/. While /v/ and /x/ are produced better by around 10-15%, the biggest difference is seen in /ɣ/. While all other tests so far have seen it produced correctly no more than 10.3% (with the exception of 19.6% for the full data), 40.5% is the highest accuracy the network has shown in producing the /ɣ/. Even though the /ɣ/ is less frequent in the plural data than in the singular, it is produced correctly nearly four times more often.

However, it appears that to produce the dependent phonemes more accurately, sacrifices were made in other areas, as nearly all other phonemes show a reduced correctness in relation to previous data. This does not include /l/ and /r/, which are fully independent and do not mutate, leading them to be produced accurately in almost all environments. The /t/ also has a high correctness, most likely because it does not show up as a mutated form in the plural at all.

While the hypothesis that mutated forms are produced less accurately seems to align with these results, it is important to note that /f/ and /g/ don't appear in mutated positions half the time as one might think from the results. Only a fifth of the mistakes with /f/ and /g/ could be ascribed to them being mutated forms. The other mistakes are most likely due to the prevalence of /Ø/, /v/, and /ɣ/, which are the mutated forms of /f/ and /g/ and show up much more in the results. Because of this, these mutated forms are possibly produced where a /f/ or /g/ would be expected, showing that with this training, there are indeed associations between the mutated forms that do not appear with other testing scenarios.

3.3.2.2 Comprehension

Once again the comprehension results seen in Table 15 show support for the hypothesis that the network cannot learn the feature it is being tested on correctly. The plural is also comprehended less well than the singular was in 3.3.1 which is in line with the pattern of better comprehension for the singular forms.

*Table 15: Percentage of correct answers for comprehension of morphosyntax in the plural*

| **Morphosyntactic feature** | | **Correctness (stdev)** |
|---|---|---|
| **Case** | Nominative | 20.6% (28.0) |
| | Genitive | 33.0% (32.6) |
| | Dative | 58.3% (34.2) |
| **Plurality** | Plural | 4.0% (23.5) |
| **Definiteness** | Indefinite | 61.1% (41.4) |
| | Definite | 43.8% (42.1) |

Unlike the data for the singular, the nominative is now comprehended significantly less well. This could indicate the problems the network had when learning the nominative are located predominantly in the plural. Once again the indefinite is produced better than the definite forms, though as determined in 3.3.1 the reason for this is unclear.

## 3.4 Definiteness

For definiteness the same numbers apply as the tests for plurality. There are 144 items in the testing data, and 432 items in the training data. The testing data now only consists of either the definite (3.4.1) or indefinite (3.4.2) forms of the odd numbered entries in Appendix A, as found in Appendix B. Each datapoint occurs in around 231 training steps.

### 3.4.1 Definite

3.4.1.1 Production

In Table 16 the results for production are shown when tested on unseen definite forms. Nearly all initial phonemes are found in the definite forms, with the exception of /h/.

*Table 16: Percentage of correct answers for the production of each phoneme in the definite*

| Phoneme | Correctness | Occurrences (of 144) | Phoneme | Correctness | Occurrences (of 144) |
|---|---|---|---|---|---|
| **Ø** | 18.0% | 1.4% | **n** | 85.4% | 4.9% |
| **p** | 95.0% | 2.1% | **$n^j$** | 82.6% | 4.9% |
| **$p^j$** | 95.7% | 2.1% | **l** | 97.2% | 4.2% |
| **b** | 55.6% | 3.5% | **$l^j$** | 97.2% | 4.2% |
| **$b^j$** | 52.6% | 3.5% | **r** | 98.5% | 4.2% |
| **f** | 57.3% | 2.8% | **$r^j$** | 96.7% | 4.2% |
| **$f^j$** | 58.3% | 2.8% | **k** | 98.0% | 2.1% |
| **v** | 2.8% | 2.8% | **$k^j$** | 98.7% | 2.1% |
| **$v^j$** | 6.5% | 2.8% | **g** | 51.8% | 3.5% |
| **m** | 61.7% | 4.9% | **$g^j$** | 53.0% | 3.5% |
| **$m^j$** | 51.3% | 4.9% | **x** | 0.0% | 0.7% |
| **t** | 62.4% | 4.9% | **$x^j$** | 1.0% | 0.7% |
| **$t^j$** | 71.7% | 4.2% | **ɣ** | 4.0% | 0.7% |
| **d** | 57.5% | 4.2% | **$ɣ^j$** | 3.0% | 0.7% |
| **$d^j$** | 51.2% | 4.2% | **ŋ** | 0.0% | 1.4% |
| **s** | 91.8% | 2.8% | **$ŋ^j$** | 0.0% | 1.4% |
| **$s^j$** | 77.2% | 3.5% | | | |

Once again the production of the dependent phonemes is distinctly lacking. Though the [$v^j$] shows a slightly higher accuracy than [v], there is no clear reason for this and the difference may simply be an arbitrary variation. All other phonemes tend to be produced correctly a similar amount of times relative to the amount that they appear in mutated and unmutated forms. The exception to this is /d/, which only occurs as an eclipsed form 1 out of 6 occurrences yet is produced correctly around 55% of the time, rather than the expected ~83%.

#### 3.4.1.2 Comprehension

The results for comprehension can be seen in Table 17. Definiteness is only comprehended correctly in 4.1%, conforming to the pattern of worse comprehension in the feature being tested on.

*Table 17: Percentage of correct answers for comprehension of morphosyntax in the definite*

| **Morphosyntactic feature** | | **Correctness (stdev)** |
|---|---|---|
| **Case** | Nominative | 24.1% (29.6) |
| | Genitive | 21.9% (28.7) |
| | Dative | 52.2% (34.6) |
| **Plurality** | Singular | 59.6% (41.6) |
| | Plural | 42.3% (41.9) |
| **Definiteness** | Definite | 4.1% (23.9) |

The data again shows that the dative is comprehended best out of all cases, and the singular is comprehended better than the plural. The difference between the singular and plural is slightly larger than in the other results, but still falls in line with the general tendency of comprehending the singular forms more accurately.

### 3.4.2 Indefinite

#### 3.4.2.1 Production

The last test was of the indefinite words. The production of these forms resulted in the outcomes seen in Table 18. Not all initial phonemes are present, as /ŋ/ is not found in the indefinite forms at all.

*Table 18: Percentage of correct answers for the production of each phoneme in the indefinite*

| Phoneme | Correctness | Occurrences (of 144) | Phoneme | Correctness | Occurrences (of 144) |
|---|---|---|---|---|---|
| Ø | 3.5% | 2.8% | $s^j$ | 99.3% | 2.8% |
| p | 71.5% | 2.8% | n | 94.8% | 4.2% |
| $p^j$ | 77.8% | 2.8% | $n^j$ | 95.7% | 4.2% |
| b | 80.3% | 2.8% | l | 96.8% | 4.2% |
| $b^j$ | 82.0% | 2.8% | $l^j$ | 95.2% | 4.2% |
| f | 30.7% | 4.2% | r | 96.5% | 4.2% |
| $f^j$ | 31.0% | 4.2% | $r^j$ | 97.5% | 4.2% |
| v | 4.8% | 2.8% | k | 76.0% | 2.8% |
| $v^j$ | 2.8% | 2.8% | $k^j$ | 76.8% | 2.8% |
| m | 96.5% | 2.8% | g | 75.0% | 2.8% |
| $m^j$ | 97.5% | 2.8% | $g^j$ | 75.0% | 2.8% |
| t | 98.8% | 2.8% | x | 0.0% | 1.4% |
| $t^j$ | 99.5% | 2.8% | $x^j$ | 0.5% | 1.4% |
| d | 99.0% | 2.8% | ɣ | 7.3% | 2.8% |
| $d^j$ | 98.3% | 2.8% | $ɣ^j$ | 7.5% | 2.8% |
| s | 94.8% | 2.8% | h | 1.8% | 5.6% |

Unlike the definite forms, the indefinite sees no eclipsis, and only few consonants lenite. In fact, except for /f/, all consonants in the indefinite data appear as an unmutated consonant only, or a mutated consonant only. The consonants that are mutated are the dependent phonemes /Ø/, /v/, /x/, /ɣ/, and /h/, while all other consonants are unmutated in the indefinite forms. Exactly this is what makes it interesting to see what independent phonemes are produced correctly, since any variation in the correctness is related to factors outside of any mutated forms present.

Some of this variation can be seen with /p/, /b/, /k/, and /g/. These are all produced incorrectly around 25% of the time, even though they show no variation in the indefinite at all. This is most likely due to the fact that these phonemes mutate in definite forms, though it is unusual that other phonemes such as /t/, /d/, /m/, and /s/ do not show this even though they also mutate a similar amount of times.

Another noticeable variation is /f/, which is produced correctly only around 30% of the time. Even though it mutates a similar amount to other phonemes, it is not treated similarly to the phonemes described above. The exact cause of this is not clear from the mutations alone.

3.4.2.2 Comprehension

Results for comprehension of the indefinite forms can be found in Table 19. As seen with the plural/singular data, the overall trend to comprehend the indefinite forms better is preserved despite the low overall correctness.

*Table 19: Percentage of correct answers for comprehension of morphosyntax in the indefinite*

| **Morphosyntactic feature** | | **Correctness (stdev)** |
|---|---|---|
| **Case** | Nominative | 28.4% (31.2) |
| | Genitive | 36.3% (33.3) |
| | Dative | 56.5% (34.3) |
| **Plurality** | Singular | 55.8% (42.1) |
| | Plural | 44.2% (42.1) |
| **Definiteness** | Indefinite | 7.1% (30.9) |

While a small difference, compared to the other data, the genitive is comprehended slightly better overall. The dative is comprehended similarly to other situations, while the nominative shows a small increase along with the genitive. In the plurality there is no specific difference between this data and other results, as the singular is comprehended correctly somewhat more than the plural is.

# 4. Comparing Phonemes

Now that the model has been tested on its performance on data, another analysis can be performed to observe any patterns the network has created to analyze the phonology. For this, the activation levels of the second layer need to be recorded after activating an individual phoneme and spreading it just like in the testing phase. After doing this for every phoneme, the activation levels can then be compared by calculating the cosine similarity for all phonemes and constructing a table with percentages. These percentages indicate the similarities between the phonemes, higher percentages indicating more similar features.

Image 3 shows the similarity matrix for all phonemes after training on the complete dataset, with 100.000 training steps, averaged over 100 learners. The colors indicate (relative) similarity, with dark green being the highest similarity, while white indicates a very low similarity (as compared to the mean percentage). Image 4 contains the standard deviations for each of the values, with dark blue indicating a lower standard deviation and white a higher standard deviation.

The first thing that stands out when looking at the data in Image 3, is how the columns for /x/ and /ŋ/ are the only columns that are completely green. This means there is much overlap between these phonemes and all other phonemes, which could explain why the model performed poorly on these phonemes in all tests. There are not enough distinct features to these phonemes for the network to know when to produce them. Even considering this, however, there is a clear increase in similarity between /x/ and /k/, versus /x/ and other phonemes. This is specifically for [x] - [k], and [$x^j$] - [$k^j$], which makes sense since /x/ is the lenited form of /k/. This same comparison can also be made between /ŋ/ and /g/, where /g/ has the highest similarity with /ŋ/ out of all phonemes. /v/, /ɣ/ and /Ø/ also show higher similarities to /x/ and /ŋ/, but this is most likely due to their status as dependent phonemes.

When looking at the standard deviations in Image 4, the dependent phonemes actually show an overall lower standard deviation in comparison with nearly all phonemes. The exceptions to this are [t], [n], [$r^j$], and /h/. For [t], [$r^j$] and /h/ this can be explained by the overall higher standard deviation, but [n] shows this pattern less than the other phonemes. Still, there seems to be a pattern where dependent phonemes show a higher similarity to other phonemes consistently.

Some of the more white-dominant columns are those of /l/ and /r/, which can be easily explained due to their lack of mutation. Since they occur for all forms of any word beginning with /l/ or /r/, they do not form large associations with any other phonemes, which leads to this relatively low score across the board. The other fully independent phonemes (/p/ /s/ /k/) do not have very low scores, as they do mutate in some situations, leading to a connection with the mutated phoneme and thus more similarity with those phonemes. Interestingly, the partially independent phoneme /n/ has a lower score when compared to the fully independent phonemes. This shows support for the hypothesis that lack of mutation causes a lower cosine similarity score, since /n/ occurs as an independent phoneme more than it does as a mutated form.

This logic then extends to the other partially independent phonemes (/f/ /b/ /m/ /t/ /d/ /g/) as well. They all occur at the start of a word in unmutated conditions but because they themselves mutate, they form connections to the phonemes they mutate into. The stronger connections arise not from the fact that they occur in the dataset as both mutated and unmutated forms, but rather the fact that they mutate *into* other phonemes that are then associated with one another.

Interestingly, /h/ is expected to pattern with /t/ and /s/ since these are the forms that can mutate into /h/, but shows similar similarities with /p/, /v/, /k/, and /g/. Setting aside the similarity with /v/ for later analysis, /p/, /k/, and /g/ should not show any stronger similarity to /h/ than other independent phonemes. While /p/ and /k/ are fully independent, /g/ is not.

Therefore there might be other phonological patterns that could be causing this increased similarity that are not related to lenition or eclipsis.

Of course the data must also be analyzed to see if lenited phonemes show a stronger connection with their unmutated counterparts compared to the eclipsed phonemes. When looking at /k/ and /g/, the /k/ shows a much stronger similarity to /x/ than to /g/, and /g/ shows a stronger similarity to /ŋ/ than it does to /ɣ/. However, when looking at /t/ and /d/, the similarity scores for /d/ are much more similar to one another, with nearly no difference in the lenited vs eclipsed forms. /t/ is somewhat more similar to /h/ than it is to /d/, but the standard deviation is also higher. Now looking at /p/ and /b/, there is a very small difference in similarity, showing that the network does not necessarily have a stronger connection between lenited-unmutated consonant pairs compared to eclipsed-unmutated pairs.

While there are many other aspects to compare, the last one that will be mentioned here is the lenition merger of /b/ and /m/ into /v/, and of /d/ and /g/ into /ɣ/. The similarities between the pairs /b/ & /v/ (~78.6%) and /m/ & /v/ (~78.6%) is almost the exact same, while the pair of /b/ & /m/ (~76.1%) is slightly less similar to the other two similarities. For /d/ & /ɣ/ (~81.6%) and /g/ & /ɣ/ (~84.3%) the similarities lie slightly higher due to /ɣ/'s general tendency towards a higher similarity but still lie quite close. For /d/ & /g/ (~75.8%) the similarity is around the same as that of /b/ & /m/. This is slightly unexpected, as /b/ eclipses into /m/, thus it would be expected to have a stronger connection than /d/ & /g/ which forms no such mutation pair.

| | O | p | pʲ | b | bʲ | f | fʲ | v | vʲ | m | mʲ | t | tʲ | d | dʲ | s | sʲ | n | nʲ | l | lʲ | r | rʲ | k | kʲ | g | gʲ | x | xʲ | ɣ | ɣʲ | ŋ | ŋʲ | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| O | 100.0 | 79.1 | 79.0 | 78.3 | 78.5 | 80.5 | 80.7 | 81.1 | 81.4 | 77.6 | 77.8 | 78.1 | 77.8 | 77.8 | 77.7 | 78.1 | 78.0 | 77.6 | 77.2 | 77.2 | 76.8 | 76.9 | 77.8 | 79.7 | 79.6 | 78.7 | 78.7 | 88.3 | 88.6 | 84.2 | 84.8 | 89.6 | 89.7 | 81.4 |
| p | 79.1 | 100.0 | 77.5 | 77.5 | 76.6 | 78.0 | 76.7 | 77.7 | 78.0 | 76.2 | 76.0 | 76.3 | 76.2 | 76.0 | 76.2 | 76.2 | 76.1 | 75.8 | 75.5 | 75.7 | 75.4 | 75.4 | 76.0 | 77.9 | 77.5 | 77.1 | 77.0 | 85.0 | 85.2 | 81.1 | 81.1 | 86.3 | 86.1 | 78.4 |
| pʲ | 79.0 | 77.5 | 100.0 | 76.4 | 77.5 | 76.4 | 78.1 | 77.6 | 77.6 | 75.8 | 75.8 | 76.5 | 76.2 | 76.0 | 76.1 | 76.5 | 76.2 | 76.0 | 75.6 | 75.9 | 75.7 | 75.7 | 76.1 | 77.7 | 77.8 | 77.2 | 77.0 | 84.8 | 85.3 | 81.2 | 81.5 | 86.2 | 86.2 | 78.6 |
| b | 78.3 | 77.5 | 76.4 | 100.0 | 75.9 | 76.0 | 75.9 | 78.3 | 77.1 | 76.0 | 75.5 | 75.4 | 75.7 | 75.5 | 75.2 | 75.9 | 75.5 | 75.4 | 75.1 | 75.4 | 74.7 | 74.8 | 75.2 | 77.0 | 76.9 | 76.5 | 76.5 | 84.3 | 84.6 | 80.4 | 80.5 | 85.4 | 85.4 | 77.8 |
| bʲ | 78.5 | 76.6 | 77.5 | 75.9 | 100.0 | 75.9 | 76.4 | 77.5 | 78.9 | 75.5 | 76.2 | 75.5 | 75.9 | 75.4 | 75.4 | 75.7 | 75.5 | 75.5 | 75.1 | 75.2 | 74.8 | 75.0 | 75.4 | 77.2 | 77.1 | 76.4 | 76.5 | 84.6 | 84.8 | 80.6 | 80.9 | 85.7 | 85.7 | 78.0 |
| f | 80.5 | 78.0 | 76.4 | 76.0 | 75.9 | 100.0 | 75.7 | 78.0 | 77.1 | 75.5 | 75.3 | 75.3 | 75.7 | 75.3 | 75.6 | 75.8 | 75.7 | 75.5 | 75.1 | 75.0 | 74.7 | 74.9 | 75.5 | 77.0 | 76.9 | 76.7 | 76.3 | 84.4 | 84.6 | 80.5 | 80.6 | 85.7 | 85.6 | 77.8 |
| fʲ | 80.7 | 76.7 | 78.1 | 75.9 | 76.4 | 75.7 | 100.0 | 77.2 | 78.1 | 75.6 | 75.5 | 75.7 | 75.8 | 75.3 | 75.4 | 76.0 | 75.5 | 75.3 | 75.1 | 75.4 | 74.9 | 74.8 | 75.6 | 77.2 | 77.1 | 76.6 | 76.4 | 84.5 | 84.9 | 80.4 | 80.6 | 85.8 | 85.8 | 77.9 |
| v | 81.1 | 77.7 | 77.6 | 78.3 | 77.5 | 78.0 | 77.2 | 100.0 | 78.9 | 78.5 | 76.6 | 76.6 | 76.5 | 76.4 | 76.6 | 76.8 | 76.7 | 76.3 | 75.9 | 76.2 | 75.9 | 75.6 | 76.4 | 78.3 | 78.1 | 77.7 | 77.7 | 86.5 | 86.8 | 82.3 | 83.0 | 87.9 | 87.9 | 79.8 |
| vʲ | 81.4 | 78.0 | 77.6 | 77.1 | 78.9 | 77.1 | 78.1 | 78.9 | 100.0 | 76.5 | 78.6 | 76.9 | 76.7 | 76.7 | 76.6 | 77.0 | 76.7 | 76.6 | 76.3 | 76.3 | 75.9 | 75.8 | 76.5 | 78.4 | 78.2 | 77.7 | 77.7 | 86.6 | 86.9 | 82.7 | 82.7 | 87.9 | 88.0 | 79.9 |
| m | 77.6 | 76.2 | 75.8 | 76.0 | 75.5 | 75.5 | 75.6 | 78.5 | 76.5 | 100.0 | 75.1 | 75.1 | 75.3 | 75.1 | 74.8 | 75.5 | 75.1 | 74.9 | 74.8 | 74.9 | 74.5 | 74.4 | 75.0 | 76.5 | 76.2 | 75.8 | 75.8 | 83.5 | 83.8 | 79.6 | 79.6 | 84.7 | 84.6 | 77.4 |
| mʲ | 77.8 | 76.0 | 75.8 | 75.5 | 76.2 | 75.3 | 75.5 | 76.6 | 78.6 | 75.1 | 100.0 | 75.4 | 75.4 | 75.1 | 75.2 | 75.2 | 75.3 | 75.1 | 74.7 | 74.7 | 74.4 | 74.6 | 75.3 | 76.4 | 76.4 | 76.2 | 75.9 | 83.8 | 84.0 | 79.8 | 80.0 | 84.8 | 84.7 | 77.4 |
| t | 78.1 | 76.3 | 76.5 | 75.4 | 75.5 | 75.3 | 75.7 | 76.6 | 76.9 | 75.1 | 75.4 | 100.0 | 75.2 | 75.3 | 75.3 | 76.6 | 75.5 | 75.2 | 75.0 | 75.1 | 74.6 | 74.8 | 75.3 | 76.8 | 76.5 | 76.2 | 75.9 | 83.8 | 84.2 | 80.0 | 80.2 | 85.2 | 85.1 | 78.6 |
| tʲ | 77.8 | 76.2 | 76.2 | 75.7 | 75.9 | 75.7 | 75.8 | 76.5 | 76.7 | 75.3 | 75.4 | 75.2 | 100.0 | 75.3 | 75.6 | 75.6 | 75.8 | 74.9 | 74.9 | 75.0 | 74.6 | 74.5 | 75.2 | 76.9 | 76.7 | 76.2 | 76.0 | 83.7 | 84.0 | 80.0 | 80.0 | 84.9 | 84.8 | 78.4 |
| d | 77.8 | 76.0 | 76.0 | 75.5 | 75.4 | 75.3 | 75.3 | 76.4 | 76.7 | 75.1 | 75.1 | 75.3 | 75.3 | 100.0 | 74.9 | 75.4 | 75.2 | 75.5 | 74.8 | 74.9 | 74.6 | 74.6 | 75.1 | 76.6 | 76.4 | 75.7 | 76.0 | 83.5 | 83.7 | 81.4 | 79.8 | 84.6 | 84.6 | 77.1 |
| dʲ | 77.7 | 76.2 | 76.1 | 75.2 | 75.4 | 75.6 | 75.4 | 76.6 | 76.6 | 74.8 | 75.2 | 75.3 | 75.6 | 74.9 | 100.0 | 75.4 | 75.2 | 75.1 | 75.0 | 74.8 | 74.4 | 74.3 | 75.2 | 76.5 | 76.5 | 76.1 | 75.9 | 83.4 | 83.8 | 79.8 | 81.8 | 84.7 | 84.6 | 77.1 |
| s | 78.1 | 76.2 | 76.5 | 75.9 | 75.7 | 75.8 | 76.0 | 76.8 | 77.0 | 75.5 | 75.2 | 76.6 | 75.6 | 75.4 | 75.4 | 100.0 | 75.7 | 75.3 | 74.9 | 75.1 | 74.8 | 74.7 | 75.4 | 77.0 | 76.9 | 76.1 | 76.2 | 84.1 | 84.1 | 80.2 | 80.4 | 85.0 | 85.0 | 78.6 |
| sʲ | 78.0 | 76.1 | 76.2 | 75.5 | 75.5 | 75.7 | 75.5 | 76.7 | 76.7 | 75.1 | 75.3 | 75.5 | 75.8 | 75.2 | 75.2 | 75.7 | 100.0 | 74.8 | 74.9 | 75.0 | 74.4 | 74.8 | 75.3 | 76.6 | 76.5 | 76.1 | 75.9 | 83.6 | 83.8 | 80.0 | 80.0 | 84.6 | 84.5 | 78.4 |
| n | 77.6 | 75.8 | 76.0 | 75.4 | 75.5 | 75.5 | 75.3 | 76.3 | 76.6 | 74.9 | 75.1 | 75.2 | 74.9 | 75.5 | 75.1 | 75.3 | 74.8 | 100.0 | 74.8 | 74.8 | 74.5 | 74.7 | 75.1 | 76.3 | 76.0 | 75.9 | 75.6 | 83.3 | 83.7 | 79.7 | 79.7 | 84.6 | 84.5 | 77.2 |
| nʲ | 77.2 | 75.5 | 75.6 | 75.1 | 75.1 | 75.1 | 75.1 | 75.9 | 76.3 | 74.8 | 74.7 | 75.0 | 74.9 | 74.8 | 75.0 | 74.9 | 74.9 | 74.8 | 100.0 | 74.7 | 74.3 | 74.4 | 74.9 | 76.0 | 75.9 | 75.5 | 75.6 | 83.0 | 83.2 | 79.2 | 79.5 | 84.1 | 84.1 | 76.6 |
| l | 77.2 | 75.7 | 75.9 | 75.4 | 75.2 | 75.0 | 75.4 | 76.2 | 76.3 | 74.9 | 74.7 | 75.1 | 75.0 | 74.9 | 74.8 | 75.1 | 75.0 | 74.8 | 74.7 | 100.0 | 74.4 | 74.4 | 75.0 | 76.1 | 75.8 | 75.6 | 75.4 | 82.9 | 83.2 | 79.4 | 79.4 | 84.1 | 84.0 | 76.7 |
| lʲ | 76.8 | 75.4 | 75.7 | 74.7 | 74.8 | 74.7 | 74.9 | 75.9 | 75.9 | 74.5 | 74.4 | 74.6 | 74.6 | 74.6 | 74.4 | 74.8 | 74.4 | 74.5 | 74.3 | 74.4 | 100.0 | 74.1 | 74.7 | 75.7 | 75.7 | 75.4 | 75.3 | 82.4 | 82.7 | 78.7 | 79.0 | 83.5 | 83.4 | 76.4 |
| r | 76.9 | 75.4 | 75.7 | 74.8 | 75.0 | 74.9 | 74.8 | 75.6 | 75.8 | 74.4 | 74.6 | 74.8 | 74.5 | 74.6 | 74.3 | 74.7 | 74.8 | 74.7 | 74.4 | 74.4 | 74.1 | 100.0 | 74.5 | 75.7 | 75.9 | 75.3 | 75.1 | 82.5 | 82.7 | 79.1 | 79.1 | 83.5 | 83.5 | 76.3 |
| rʲ | 77.8 | 76.0 | 76.1 | 75.2 | 75.4 | 75.5 | 75.6 | 76.4 | 76.5 | 75.0 | 75.3 | 75.3 | 75.2 | 75.1 | 75.2 | 75.4 | 75.3 | 75.1 | 74.9 | 75.0 | 74.7 | 74.5 | 100.0 | 76.5 | 76.3 | 75.8 | 75.9 | 83.3 | 83.6 | 79.5 | 79.7 | 84.5 | 84.4 | 77.3 |
| k | 79.7 | 77.9 | 77.7 | 77.0 | 77.2 | 77.0 | 77.2 | 78.3 | 78.4 | 76.5 | 76.4 | 76.8 | 76.9 | 76.6 | 76.5 | 77.0 | 76.6 | 76.3 | 76.0 | 76.1 | 75.7 | 75.7 | 76.5 | 100.0 | 78.3 | 78.8 | 77.5 | 90.6 | 86.0 | 81.7 | 81.9 | 86.8 | 86.9 | 78.9 |
| kʲ | 79.6 | 77.5 | 77.8 | 76.9 | 77.1 | 76.9 | 77.1 | 78.1 | 78.2 | 76.2 | 76.4 | 76.5 | 76.7 | 76.4 | 76.5 | 76.9 | 76.5 | 76.0 | 75.9 | 75.8 | 75.7 | 75.9 | 76.3 | 78.3 | 100.0 | 77.7 | 78.4 | 85.5 | 90.3 | 81.5 | 81.5 | 86.7 | 86.5 | 78.8 |
| g | 78.7 | 77.1 | 77.2 | 76.5 | 76.4 | 76.7 | 76.6 | 77.7 | 77.7 | 75.8 | 76.2 | 76.2 | 76.2 | 75.7 | 76.1 | 76.1 | 76.1 | 75.9 | 75.5 | 75.6 | 75.4 | 75.3 | 75.8 | 78.8 | 77.7 | 100.0 | 76.9 | 85.9 | 85.3 | 84.2 | 81.2 | 88.2 | 86.3 | 78.5 |
| gʲ | 78.7 | 77.0 | 77.0 | 76.5 | 76.5 | 76.3 | 76.4 | 77.7 | 77.7 | 75.8 | 75.9 | 75.9 | 76.0 | 76.0 | 75.9 | 76.2 | 75.9 | 75.6 | 75.6 | 75.4 | 75.3 | 75.1 | 75.9 | 77.5 | 78.4 | 76.9 | 100.0 | 84.9 | 86.1 | 80.9 | 84.3 | 86.3 | 88.1 | 78.4 |
| x | 88.3 | 85.0 | 84.8 | 84.3 | 84.6 | 84.4 | 84.5 | 86.5 | 86.6 | 83.5 | 83.8 | 83.8 | 83.7 | 83.5 | 83.4 | 84.1 | 83.6 | 83.3 | 83.0 | 82.9 | 82.4 | 82.5 | 83.3 | 90.6 | 85.5 | 85.9 | 84.9 | 100.0 | 96.2 | 91.3 | 91.6 | 97.2 | 97.2 | 87.7 |
| xʲ | 88.6 | 85.2 | 85.3 | 84.6 | 84.8 | 84.6 | 84.9 | 86.8 | 86.9 | 83.8 | 84.0 | 84.2 | 84.0 | 83.7 | 83.8 | 84.1 | 83.8 | 83.7 | 83.2 | 83.2 | 82.7 | 82.7 | 83.6 | 86.0 | 90.3 | 85.3 | 86.1 | 96.2 | 100.0 | 91.6 | 91.5 | 97.6 | 97.4 | 88.0 |
| ɣ | 84.2 | 81.1 | 81.2 | 80.4 | 80.6 | 80.5 | 80.4 | 82.3 | 82.7 | 79.6 | 79.8 | 80.0 | 80.0 | 81.4 | 79.8 | 80.2 | 80.0 | 79.7 | 79.2 | 79.4 | 78.7 | 79.1 | 79.5 | 81.7 | 81.5 | 84.2 | 80.9 | 91.3 | 91.6 | 100.0 | 87.5 | 94.8 | 92.4 | 83.8 |
| ɣʲ | 84.8 | 81.1 | 81.5 | 80.5 | 80.9 | 80.6 | 80.6 | 83.0 | 82.7 | 79.6 | 80.0 | 80.2 | 80.0 | 79.8 | 81.8 | 80.4 | 80.0 | 79.7 | 79.5 | 79.4 | 79.0 | 79.1 | 79.7 | 81.9 | 81.5 | 81.2 | 84.3 | 91.6 | 91.5 | 87.5 | 100.0 | 92.8 | 95.0 | 84.0 |
| ŋ | 89.6 | 86.3 | 86.2 | 85.4 | 85.7 | 85.7 | 85.8 | 87.9 | 87.9 | 84.7 | 84.8 | 85.2 | 84.9 | 84.6 | 84.7 | 85.0 | 84.6 | 84.6 | 84.1 | 84.1 | 83.5 | 83.5 | 84.5 | 86.8 | 86.7 | 88.2 | 86.3 | 97.2 | 97.6 | 94.8 | 92.8 | 100.0 | 98.9 | 89.1 |
| ŋʲ | 89.7 | 86.1 | 86.2 | 85.4 | 85.7 | 85.6 | 85.8 | 87.9 | 88.0 | 84.6 | 84.7 | 85.1 | 84.8 | 84.6 | 84.6 | 85.0 | 84.5 | 84.5 | 84.1 | 84.0 | 83.4 | 83.5 | 84.4 | 86.9 | 86.5 | 86.3 | 88.1 | 97.2 | 97.4 | 92.4 | 95.0 | 98.9 | 100.0 | 89.1 |
| h | 81.4 | 78.4 | 78.6 | 77.8 | 78.0 | 77.8 | 77.9 | 79.8 | 79.9 | 77.4 | 77.4 | 78.6 | 78.4 | 77.1 | 77.1 | 78.6 | 78.4 | 77.2 | 76.6 | 76.7 | 76.4 | 76.3 | 77.3 | 78.9 | 78.8 | 78.5 | 78.4 | 87.7 | 88.0 | 83.8 | 84.0 | 89.1 | 89.1 | 100.0 |

*Image 3: Cosine similarity matrix of the initial phonemes in Irish of 100 learners*

| | O | p | pʲ | b | bʲ | f | fʲ | v | vʲ | m | mʲ | t | tʲ | d | dʲ | s | sʲ | n | nʲ | l | lʲ | r | rʲ | k | kʲ | g | gʲ | x | xʲ | ɣ | ɣʲ | ŋ | ŋʲ | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| O | 0.00 | 2.12 | 2.36 | 2.38 | 2.26 | 2.44 | 2.41 | 2.62 | 2.55 | 2.69 | 2.67 | 2.82 | 2.43 | 2.45 | 2.41 | 2.39 | 2.53 | 2.72 | 2.76 | 2.80 | 2.49 | 2.82 | 2.90 | 2.34 | 2.24 | 2.43 | 2.50 | 2.27 | 2.03 | 2.56 | 2.73 | 1.91 | 2.03 | 2.51 |
| p | 2.12 | 0.00 | 2.25 | 2.43 | 2.44 | 2.20 | 2.40 | 2.55 | 2.29 | 2.36 | 2.43 | 2.67 | 2.53 | 2.31 | 2.72 | 2.34 | 2.51 | 2.60 | 2.43 | 2.67 | 2.42 | 2.83 | 2.71 | 2.14 | 2.49 | 2.29 | 2.36 | 1.99 | 2.10 | 2.59 | 2.42 | 1.84 | 1.84 | 2.51 |
| pʲ | 2.36 | 2.25 | 0.00 | 2.47 | 2.19 | 2.10 | 2.11 | 2.41 | 2.30 | 2.56 | 2.35 | 2.77 | 2.15 | 2.16 | 2.21 | 2.16 | 2.22 | 2.63 | 2.42 | 2.53 | 2.40 | 2.68 | 2.32 | 2.04 | 2.10 | 2.00 | 2.06 | 1.96 | 1.84 | 2.52 | 2.51 | 1.60 | 1.63 | 2.27 |
| b | 2.38 | 2.43 | 2.47 | 0.00 | 2.60 | 2.44 | 2.34 | 2.65 | 2.48 | 2.69 | 2.38 | 3.12 | 2.64 | 2.38 | 2.39 | 2.48 | 2.22 | 2.76 | 2.60 | 2.49 | 2.41 | 2.75 | 2.69 | 2.15 | 2.34 | 2.31 | 2.34 | 1.94 | 1.93 | 2.37 | 2.39 | 1.80 | 1.79 | 2.61 |
| bʲ | 2.26 | 2.44 | 2.19 | 2.60 | 0.00 | 2.66 | 2.28 | 2.45 | 2.49 | 2.51 | 2.61 | 2.76 | 2.50 | 2.50 | 2.13 | 2.16 | 2.42 | 2.56 | 2.29 | 2.76 | 2.52 | 2.66 | 2.71 | 2.28 | 2.34 | 2.31 | 2.39 | 2.08 | 2.05 | 2.21 | 2.50 | 1.74 | 1.80 | 2.47 |
| f | 2.44 | 2.20 | 2.10 | 2.44 | 2.66 | 0.00 | 2.45 | 2.60 | 2.52 | 2.53 | 2.50 | 2.89 | 2.56 | 2.37 | 2.48 | 2.23 | 2.31 | 2.41 | 2.50 | 2.69 | 2.41 | 2.93 | 2.60 | 2.23 | 2.25 | 2.22 | 2.40 | 1.98 | 2.16 | 2.44 | 2.42 | 1.82 | 1.96 | 2.74 |
| fʲ | 2.41 | 2.40 | 2.11 | 2.34 | 2.28 | 2.45 | 0.00 | 2.73 | 2.41 | 2.73 | 2.67 | 2.84 | 2.56 | 2.27 | 2.40 | 2.49 | 2.23 | 2.54 | 2.63 | 2.72 | 2.61 | 2.97 | 2.72 | 2.20 | 2.39 | 2.36 | 2.27 | 2.07 | 2.19 | 2.23 | 2.32 | 1.85 | 2.01 | 2.44 |
| v | 2.62 | 2.55 | 2.41 | 2.65 | 2.45 | 2.60 | 2.73 | 0.00 | 2.39 | 2.64 | 2.64 | 3.19 | 2.57 | 2.28 | 2.41 | 2.44 | 2.43 | 2.77 | 2.57 | 2.82 | 2.64 | 2.97 | 2.83 | 2.21 | 2.42 | 2.50 | 2.39 | 2.20 | 2.07 | 2.55 | 2.67 | 1.75 | 1.88 | 2.65 |
| vʲ | 2.55 | 2.29 | 2.30 | 2.48 | 2.49 | 2.52 | 2.41 | 2.39 | 0.00 | 2.60 | 2.41 | 2.71 | 2.65 | 2.43 | 2.62 | 2.32 | 2.31 | 2.88 | 2.50 | 2.89 | 2.73 | 2.65 | 2.69 | 2.16 | 2.11 | 2.20 | 2.35 | 1.92 | 1.97 | 2.58 | 2.34 | 1.83 | 1.85 | 2.72 |
| m | 2.69 | 2.36 | 2.56 | 2.69 | 2.51 | 2.53 | 2.73 | 2.64 | 2.60 | 0.00 | 2.60 | 2.90 | 2.62 | 2.25 | 2.63 | 2.24 | 2.51 | 2.73 | 2.50 | 2.97 | 2.58 | 2.78 | 2.53 | 2.28 | 2.57 | 2.41 | 2.39 | 2.47 | 2.42 | 2.51 | 2.62 | 2.16 | 2.29 | 2.79 |
| mʲ | 2.67 | 2.43 | 2.35 | 2.38 | 2.61 | 2.50 | 2.67 | 2.64 | 2.41 | 2.60 | 0.00 | 2.89 | 2.51 | 2.53 | 2.58 | 2.22 | 2.48 | 2.70 | 2.65 | 2.50 | 2.44 | 2.90 | 2.68 | 2.24 | 2.57 | 2.42 | 2.50 | 2.13 | 2.32 | 2.57 | 2.74 | 2.26 | 2.27 | 2.78 |
| t | 2.82 | 2.67 | 2.77 | 3.12 | 2.76 | 2.89 | 2.84 | 3.19 | 2.71 | 2.90 | 2.89 | 0.00 | 2.78 | 2.84 | 2.74 | 2.62 | 2.63 | 2.89 | 2.84 | 2.85 | 3.00 | 2.92 | 2.87 | 2.62 | 2.72 | 2.67 | 2.65 | 2.79 | 2.89 | 2.90 | 2.86 | 2.64 | 2.70 | 2.78 |
| tʲ | 2.43 | 2.53 | 2.15 | 2.64 | 2.50 | 2.56 | 2.56 | 2.57 | 2.65 | 2.62 | 2.51 | 2.78 | 0.00 | 2.57 | 2.37 | 2.41 | 2.38 | 2.81 | 2.53 | 2.59 | 2.70 | 2.64 | 2.87 | 2.31 | 2.21 | 2.49 | 2.33 | 2.34 | 2.39 | 2.49 | 2.45 | 2.10 | 2.05 | 2.68 |
| d | 2.45 | 2.31 | 2.16 | 2.38 | 2.50 | 2.37 | 2.27 | 2.28 | 2.43 | 2.25 | 2.53 | 2.84 | 2.57 | 0.00 | 2.27 | 2.34 | 2.29 | 2.69 | 2.35 | 2.46 | 2.33 | 2.67 | 2.41 | 2.09 | 2.28 | 2.18 | 2.13 | 1.96 | 1.92 | 2.52 | 2.45 | 1.80 | 1.80 | 2.42 |
| dʲ | 2.41 | 2.72 | 2.21 | 2.39 | 2.13 | 2.48 | 2.40 | 2.41 | 2.62 | 2.63 | 2.58 | 2.74 | 2.37 | 2.27 | 0.00 | 2.12 | 2.21 | 2.49 | 2.39 | 2.60 | 2.46 | 2.52 | 2.39 | 2.15 | 2.05 | 2.38 | 2.25 | 2.01 | 1.93 | 2.38 | 2.49 | 1.68 | 1.75 | 2.65 |
| s | 2.39 | 2.34 | 2.16 | 2.48 | 2.16 | 2.23 | 2.49 | 2.44 | 2.32 | 2.24 | 2.22 | 2.62 | 2.41 | 2.34 | 2.12 | 0.00 | 2.22 | 2.55 | 2.54 | 2.58 | 2.39 | 2.61 | 2.32 | 2.29 | 2.50 | 2.16 | 2.13 | 1.88 | 1.99 | 2.42 | 2.55 | 1.83 | 1.86 | 2.34 |
| sʲ | 2.53 | 2.51 | 2.22 | 2.22 | 2.42 | 2.31 | 2.23 | 2.43 | 2.31 | 2.51 | 2.48 | 2.63 | 2.38 | 2.29 | 2.21 | 2.22 | 0.00 | 2.48 | 2.42 | 2.49 | 2.44 | 2.62 | 2.45 | 1.96 | 2.37 | 2.19 | 2.06 | 1.88 | 1.95 | 2.37 | 2.40 | 1.74 | 1.78 | 2.46 |
| n | 2.72 | 2.60 | 2.63 | 2.76 | 2.56 | 2.41 | 2.54 | 2.77 | 2.88 | 2.73 | 2.70 | 2.89 | 2.81 | 2.69 | 2.49 | 2.55 | 2.48 | 0.00 | 2.57 | 2.77 | 2.62 | 2.88 | 2.93 | 2.59 | 2.46 | 2.58 | 2.39 | 2.98 | 2.80 | 2.98 | 2.88 | 2.69 | 2.69 | 2.85 |
| nʲ | 2.76 | 2.43 | 2.42 | 2.60 | 2.29 | 2.50 | 2.63 | 2.57 | 2.50 | 2.50 | 2.65 | 2.84 | 2.53 | 2.35 | 2.39 | 2.54 | 2.42 | 2.57 | 0.00 | 2.69 | 2.61 | 2.81 | 2.84 | 2.24 | 2.54 | 2.57 | 2.22 | 2.47 | 2.25 | 2.76 | 2.80 | 2.39 | 2.37 | 2.92 |
| l | 2.80 | 2.67 | 2.53 | 2.49 | 2.76 | 2.69 | 2.72 | 2.82 | 2.89 | 2.97 | 2.50 | 2.85 | 2.59 | 2.46 | 2.60 | 2.58 | 2.49 | 2.77 | 2.69 | 0.00 | 2.73 | 2.92 | 2.76 | 2.51 | 2.52 | 2.50 | 2.47 | 2.65 | 2.38 | 2.58 | 2.90 | 2.43 | 2.47 | 3.12 |
| lʲ | 2.49 | 2.42 | 2.40 | 2.41 | 2.52 | 2.41 | 2.61 | 2.64 | 2.73 | 2.58 | 2.44 | 3.00 | 2.70 | 2.33 | 2.46 | 2.39 | 2.44 | 2.62 | 2.61 | 2.73 | 0.00 | 2.80 | 2.70 | 2.32 | 2.63 | 2.42 | 2.43 | 2.37 | 2.43 | 2.69 | 2.87 | 2.35 | 2.41 | 2.54 |
| r | 2.82 | 2.83 | 2.68 | 2.75 | 2.66 | 2.93 | 2.97 | 2.97 | 2.65 | 2.78 | 2.90 | 2.92 | 2.64 | 2.67 | 2.52 | 2.61 | 2.62 | 2.88 | 2.81 | 2.92 | 2.80 | 0.00 | 2.53 | 2.69 | 2.81 | 2.74 | 2.56 | 2.68 | 2.48 | 3.01 | 2.84 | 2.57 | 2.59 | 2.95 |
| rʲ | 2.90 | 2.71 | 2.32 | 2.69 | 2.71 | 2.60 | 2.72 | 2.83 | 2.69 | 2.53 | 2.68 | 2.87 | 2.87 | 2.41 | 2.39 | 2.32 | 2.45 | 2.93 | 2.84 | 2.76 | 2.70 | 2.53 | 0.00 | 2.37 | 2.80 | 2.59 | 2.59 | 2.86 | 2.87 | 3.07 | 3.26 | 2.68 | 2.79 | 3.14 |
| k | 2.34 | 2.14 | 2.04 | 2.15 | 2.28 | 2.23 | 2.20 | 2.21 | 2.16 | 2.28 | 2.24 | 2.62 | 2.31 | 2.09 | 2.15 | 2.29 | 1.96 | 2.59 | 2.24 | 2.51 | 2.32 | 2.69 | 2.37 | 0.00 | 2.11 | 1.95 | 1.99 | 1.60 | 1.64 | 2.07 | 2.25 | 1.51 | 1.47 | 2.30 |
| kʲ | 2.24 | 2.49 | 2.10 | 2.34 | 2.34 | 2.25 | 2.39 | 2.42 | 2.11 | 2.57 | 2.57 | 2.72 | 2.21 | 2.28 | 2.05 | 2.50 | 2.37 | 2.46 | 2.54 | 2.52 | 2.63 | 2.81 | 2.80 | 2.11 | 0.00 | 2.17 | 2.22 | 1.92 | 1.76 | 2.31 | 2.24 | 1.67 | 1.74 | 2.54 |
| g | 2.43 | 2.29 | 2.00 | 2.31 | 2.31 | 2.22 | 2.36 | 2.50 | 2.20 | 2.41 | 2.42 | 2.67 | 2.49 | 2.18 | 2.38 | 2.16 | 2.19 | 2.58 | 2.57 | 2.50 | 2.42 | 2.74 | 2.59 | 1.95 | 2.17 | 0.00 | 2.25 | 1.95 | 1.91 | 2.77 | 2.41 | 1.75 | 1.79 | 2.65 |
| gʲ | 2.50 | 2.36 | 2.06 | 2.34 | 2.39 | 2.40 | 2.27 | 2.39 | 2.35 | 2.39 | 2.50 | 2.65 | 2.33 | 2.13 | 2.25 | 2.13 | 2.06 | 2.39 | 2.22 | 2.47 | 2.43 | 2.56 | 2.59 | 1.99 | 2.22 | 2.25 | 0.00 | 1.77 | 1.83 | 2.07 | 2.62 | 1.57 | 1.58 | 2.46 |
| x | 2.27 | 1.99 | 1.96 | 1.94 | 2.08 | 1.98 | 2.07 | 2.20 | 1.92 | 2.47 | 2.13 | 2.79 | 2.34 | 1.96 | 2.01 | 1.88 | 1.88 | 2.98 | 2.47 | 2.65 | 2.37 | 2.68 | 2.86 | 1.60 | 1.92 | 1.95 | 1.77 | 0.00 | 1.35 | 1.85 | 2.00 | 1.00 | 1.06 | 2.57 |
| xʲ | 2.03 | 2.10 | 1.84 | 1.93 | 2.05 | 2.16 | 2.19 | 2.07 | 1.97 | 2.42 | 2.32 | 2.89 | 2.39 | 1.92 | 1.93 | 1.99 | 1.95 | 2.80 | 2.25 | 2.38 | 2.43 | 2.48 | 2.87 | 1.64 | 1.76 | 1.91 | 1.83 | 1.35 | 0.00 | 1.83 | 1.97 | 0.85 | 1.05 | 2.17 |
| ɣ | 2.56 | 2.59 | 2.52 | 2.37 | 2.21 | 2.44 | 2.23 | 2.55 | 2.58 | 2.51 | 2.57 | 2.90 | 2.49 | 2.52 | 2.38 | 2.42 | 2.37 | 2.98 | 2.76 | 2.58 | 2.69 | 3.01 | 3.07 | 2.07 | 2.31 | 2.77 | 2.07 | 1.85 | 1.83 | 0.00 | 2.52 | 1.73 | 1.78 | 2.65 |
| ɣʲ | 2.73 | 2.42 | 2.51 | 2.39 | 2.50 | 2.42 | 2.32 | 2.67 | 2.34 | 2.62 | 2.74 | 2.86 | 2.45 | 2.45 | 2.49 | 2.55 | 2.40 | 2.88 | 2.80 | 2.90 | 2.87 | 2.84 | 3.26 | 2.25 | 2.24 | 2.41 | 2.62 | 2.00 | 1.97 | 2.52 | 0.00 | 1.91 | 1.78 | 2.90 |
| ŋ | 1.91 | 1.84 | 1.60 | 1.80 | 1.74 | 1.82 | 1.85 | 1.75 | 1.83 | 2.16 | 2.26 | 2.64 | 2.10 | 1.80 | 1.68 | 1.83 | 1.74 | 2.69 | 2.39 | 2.43 | 2.35 | 2.57 | 2.68 | 1.51 | 1.67 | 1.75 | 1.57 | 1.00 | 0.85 | 1.73 | 1.91 | 0.00 | 0.47 | 2.21 |
| ŋʲ | 2.03 | 1.84 | 1.63 | 1.79 | 1.80 | 1.96 | 2.01 | 1.88 | 1.85 | 2.29 | 2.27 | 2.70 | 2.05 | 1.80 | 1.75 | 1.86 | 1.78 | 2.69 | 2.37 | 2.47 | 2.41 | 2.59 | 2.79 | 1.47 | 1.74 | 1.79 | 1.58 | 1.06 | 1.05 | 1.78 | 1.78 | 0.47 | 0.00 | 2.25 |
| h | 2.51 | 2.51 | 2.27 | 2.61 | 2.47 | 2.74 | 2.44 | 2.65 | 2.72 | 2.79 | 2.78 | 2.78 | 2.68 | 2.42 | 2.65 | 2.34 | 2.46 | 2.85 | 2.92 | 3.12 | 2.54 | 2.95 | 3.14 | 2.30 | 2.54 | 2.65 | 2.46 | 2.57 | 2.17 | 2.65 | 2.90 | 2.21 | 2.25 | 0.00 |

*Image 4: Standard deviations of the cosine matrix in Image 3*

# 5. Discussion

Though the goal of this paper was to see if a Neural Network could learn the Irish IMs, this model was quite simple in design. Because of time and knowledge constraints on the author's part, there are several aspects that were not looked at during the process of creating the model. One of these aspects was variation of the model's **layering** and the possibility of adding an extra layer to combine lexical nodes to phoneme nodes before connecting to the morphosyntactic nodes, and adding an extra layer to combine the morphosyntax nodes to the lexical nodes before connecting to the phoneme nodes.

Another limitation of the current study was the **data**. While each initial phoneme was represented in the dataset, there were only two lexical entries per phoneme, which is not many considering the number of lexemes in existence. Irish also makes use of gender for nouns, which influences the mutations' occurrences. While this was accounted for in the lexeme selection (every initial phoneme was represented by two words of the same gender, so the mutations would occur consistently between the two datapoints), it would have been better to either choose words of only one gender for the entire dataset, or include gender as a morphosyntactic feature in any future models so that it can also be analyzed. This feature could cause patterns to be more accurately analyzed in relation to the real-world processing of language.

Though in 1.1 a separation was made between **fully independent phonemes and partially independent phonemes**, these groups did not hold up in relation to the results. A more accurate grouping would be to separate initial phonemes by ability to mutate, as all results were mainly influenced by the mutations (or lack thereof). However, this was not the only factor in correctness, as /f/ was shown to be produced much less reliably than other consonants that appear in similar conditions. On the other hand, out of the dependent phonemes, /ɣ/ showed the most correct production even though there was no clear cause for this either, indicating that there are more factors than just mutation that play a role in the production.

What was also handled somewhat inaccurately, was the **comprehension**. As mentioned in 3.1.2, comprehension is based purely on the lexical entry and the initial phoneme. While one could argue this is to make the Network focus on the mutations, the lack of plurality indicators (i.e. palatalization, suffixes) skews the results to be less accurate for that feature, even though most real listeners would be able to determine plurality from these indicators only, making the mutation secondary in determining the plurality. This could prevent some of the mistakes that the Network made, though this is naturally not certain until further research is done.

In comprehension, there was also a clear pattern of inability to reproduce any feature that was partially left out during **testing**. The network had a tendency to produce the dative, singular, and indefinite forms better overall, which could either indicate these forms are easier to acquire in comprehension, or the network makes a generalization towards those groups when it should not. Although mutations were first thought to cause these forms to be less variable, this turned out not to be accurate and can therefore not be an explanation as to why these forms were comprehended more accurately.

The phoneme **cosine similarity matrix** also showed various patterns that resembled the mutations and their associated connections between phonemes. With further analysis these patterns could also uncover the patterns of production and comprehension, showing why the network made certain mistakes.

Finally, certain **variations** in the mutations were purposefully ignored here, but could prove to be important in acquiring certain mutations. The /s/ → /t/ mutation, which was included in the dataset but not elaborated on in the analyses, appeared to have an influence on the results nonetheless. The allophonic variation of [w, v] and [ɣ$^j$, j] could have an influence on the network's perception of the mergers leading to those mutations and on the further distinctiveness of palatalized consonants. However, for this the dataset would need to be much larger and narrow transcriptions would be required to allow such analysis to occur.

All of these issues together influenced the ability to answer the research question. If the neural network was able to learn the patterns of the mutations, it would be possible to draw further **conclusions** from the data. Those conclusions could then help support or disprove the hypothesis that there are indeed still phonological patterns in Irish initial mutations. Since the network did not actually acquire the mutations and performed much worse on the Irish data than the network of Boersma, Chládková & Benders (2021) did on the toy language data, it is not possible to determine if any phonological patterns are present in the mutations themselves.

# 6. Conclusion

After analyzing all the data, it is clear that there is indeed a strong distinction between dependent and independent phonemes. The network had a very hard time acquiring the dependent phonemes, while the fully independent phonemes were produced and comprehended much more consistently. While the network was able to learn most of the common patterns, both lenition and eclipsis were difficult to produce regardless of the phonemes involved.

However, when the consideration must be made on whether or not this Neural Network was able to acquire Irish in production and comprehension, the answer must be no. It produced many forms, but even though it was trained thoroughly it could not consistently produce mutated forms. While comprehension was acceptable in some forms, the large amount of mistakes was too significant to consider the forms acquired.

Furthermore, to truly be able to test the main research question of whether or not Irish initial mutations are phonological, morphological, or both, a different network would be required that is actually able to learn these mutations, because without a network capable of learning the mutations, it is not possible to truly analyze any patterns that arise in either production or comprehension.

Perhaps a Neural Network with more nodes, layers, or with different functions could acquire the intricate forms of Irish initial mutations, but as it stands, this virtual learner is simply incapable.

# Bibliography

Boersma, Paul, Kateřina Chládková & Titia Benders. 2021. Phonological features emerge substance-freely from the phonetics and the morphology.

Foras na Gaeilge. 2022. Teanglann: Dictionary and language library. https://www.teanglann.ie/en/ (8 April 2022, 18 April 2022, 24 May 2022).

Green, Antony D. 2006. The independence of phonology and morphology: The Celtic Mutations. *Lingua* 116. 1946-1985.

Hickey, Richard. 2014. *The sound structure of modern Irish*. Berlin Boston: De Gruyter Mouton.

Meijer, Sandra. 2021. The emergence of consonant categories in a virtual brain. Amsterdam: University of Amsterdam BA thesis.

Terlou, Rorick. 2021. Phonetic category emergence and adaptation in a bidirectional phonological and phonetic neural network. Amsterdam: University of Amsterdam BA thesis.

# Appendix A

*Table A.1: The numbers of each phoneme node*

| Node Number | Initial Phoneme | Node Number | Initial Phoneme |
|---|---|---|---|
| 0 | Ø | 17 | n |
| 1 | p | 18 | $n^j$ |
| 2 | $p^j$ | 19 | l |
| 3 | b | 20 | $l^j$ |
| 4 | $b^j$ | 21 | r |
| 5 | f | 22 | $r^j$ |
| 6 | $f^j$ | 23 | k |
| 7 | v | 24 | $k^j$ |
| 8 | $v^j$ | 25 | g |
| 9 | m | 26 | $g^j$ |
| 10 | $m^j$ | 27 | x |
| 11 | t | 28 | $x^j$ |
| 12 | $t^j$ | 29 | ɣ |
| 13 | d | 30 | $ɣ^j$ |
| 14 | $d^j$ | 31 | ŋ |
| 15 | s | 32 | $ŋ^j$ |
| 16 | $s^j$ | 33 | h |

*Table A.2: The node numbers of each lexical node*

| Node Number | Word | Translation | Node Number | Word | Translation |
|---|---|---|---|---|---|
| 34 | Mála | Bag | 58 | Súil | Eye |
| 35 | Madra | Dog | 59 | Samhain | November |
| 36 | Méadú | Increase | 60 | Sionnach | Fox |
| 37 | Meabhlaire | Deceiver | 61 | Siopa | Shop |
| 38 | Bó | Cow | 62 | Cat | Cat |
| 39 | Bocóid | Stud/boss | 63 | Cara | Friend |
| 40 | Beoir | Beer | 64 | Ceo | Fog |
| 41 | Bealtaine | May (month) | 65 | Ceadúnas | License |
| 42 | Post | Job | 66 | Luch | Mouse |
| 43 | Páiste | Child | 67 | Lámh | Hand |
| 44 | Piobar | Pepper | 68 | Leabhar | Book |
| 45 | Peiteal | Petal | 69 | Lictéar | Lighter |

| | | | | | |
|---|---|---|---|---|---|
| **46** | Fuinneog | Window | 70 | Rón | Seal |
| **47** | Fadhb | Knot | 71 | Rothar | Bicycle |
| **48** | Fia | Deer | 72 | Réalta | Star |
| **49** | Féileacán | Butterfly | 73 | Ríog | Spasm |
| **50** | Dán | Poem | 74 | Gúna | Dress |
| **51** | Dalta | Student | 75 | Gamhain | Calf |
| **52** | Deán | Channel in strand | 76 | Gearán | Complaint |
| **53** | Dia | God | 77 | Gineadóir | Generator |
| **54** | Toradh | Fruit | 78 | Nuachtán | Newspaper |
| **55** | Tábhairne | Tavern | 79 | Nós | Tradition |
| **56** | Teanga | Language | 80 | Niteoir | Washer (person) |
| **57** | Tír | Country | 81 | Néarchóras | Nervous system |

*Table A.3: The node numbers of each morphosyntactic node*

| Morphosyntactic Feature | Nr |
|---|---|
| **NOM** | 82 |
| **GEN** | 83 |
| **DAT** | 84 |
| **SG** | 85 |
| **PL** | 86 |
| **INDEF** | 87 |
| **DEF** | 88 |

# Appendix B

| word | translation | initial phoneme | case | plurality | definiteness | process |
|---|---|---|---|---|---|---|
| **mála** | bag | m | NOM | SG | INDEF | No |
| **mála** | bag | m | NOM | SG | DEF | No |
| **málaí** | bag | m | NOM | PL | INDEF | No |
| **málaí** | bag | m | NOM | PL | DEF | No |
| **mála** | bag | m | GEN | SG | INDEF | No |
| **mhála** | bag | v | GEN | SG | DEF | Lenition |
| **málaí** | bag | m | GEN | PL | INDEF | No |
| **málaí** | bag | m | GEN | PL | DEF | No |
| **mhála** | bag | v | DAT | SG | INDEF | Lenition |
| **mála** | bag | m | DAT | SG | DEF | No |
| **mhálaí** | bag | v | DAT | PL | INDEF | Lenition |
| **málaí** | bag | m | DAT | PL | DEF | No |
| **beoir** | beer | $b^j$ | NOM | SG | INDEF | No |
| **bheoir** | beer | $v^j$ | NOM | SG | DEF | Lenition |
| **beoracha** | beer | $b^j$ | NOM | PL | INDEF | No |
| **beoracha** | beer | $b^j$ | NOM | PL | DEF | No |
| **beorach** | beer | $b^j$ | GEN | SG | INDEF | No |
| **beorach** | beer | $b^j$ | GEN | SG | DEF | No |
| **beoracha** | beer | $b^j$ | GEN | PL | INDEF | No |
| **mbeoracha** | beer | $m^j$ | GEN | PL | DEF | Eclipsis |
| **bheoir** | beer | $v^j$ | DAT | SG | INDEF | Lenition |
| **mbeoir** | beer | $m^j$ | DAT | SG | DEF | Eclipsis |
| **bheoracha** | beer | $v^j$ | DAT | PL | INDEF | Lenition |
| **beoracha** | beer | $b^j$ | DAT | PL | DEF | No |
| **rothar** | bicycle | r | NOM | SG | INDEF | No |
| **rothar** | bicycle | r | NOM | SG | DEF | No |
| **rothair** | bicycle | r | NOM | PL | INDEF | No |
| **rothair** | bicycle | r | NOM | PL | DEF | No |
| **rothair** | bicycle | r | GEN | SG | INDEF | No |
| **rothair** | bicycle | r | GEN | SG | DEF | No |
| **rothar** | bicycle | r | GEN | PL | INDEF | No |
| **rothar** | bicycle | r | GEN | PL | DEF | No |
| **rothar** | bicycle | r | DAT | SG | INDEF | No |
| **rothar** | bicycle | r | DAT | SG | DEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **rothair** | bicycle | r | DAT | PL | INDEF | No |
| **rothair** | bicycle | r | DAT | PL | DEF | No |
| **leabhar** | book | $l^j$ | NOM | SG | INDEF | No |
| **leabhar** | book | $l^j$ | NOM | SG | DEF | No |
| **leabhair** | book | $l^j$ | NOM | PL | INDEF | No |
| **leabhair** | book | $l^j$ | NOM | PL | DEF | No |
| **leabhair** | book | $l^j$ | GEN | SG | INDEF | No |
| **leabhair** | book | $l^j$ | GEN | SG | DEF | No |
| **leabhar** | book | $l^j$ | GEN | PL | INDEF | No |
| **leabhar** | book | $l^j$ | GEN | PL | DEF | No |
| **leabhar** | book | $l^j$ | DAT | SG | INDEF | No |
| **leabhar** | book | $l^j$ | DAT | SG | DEF | No |
| **leabhair** | book | $l^j$ | DAT | PL | INDEF | No |
| **leabhair** | book | $l^j$ | DAT | PL | DEF | No |
| **féileacán** | butterfly | $f^j$ | NOM | SG | INDEF | No |
| **féileacán** | butterfly | $f^j$ | NOM | SG | DEF | No |
| **féileacáin** | butterfly | $f^j$ | NOM | PL | INDEF | No |
| **féileacáin** | butterfly | $f^j$ | NOM | PL | DEF | No |
| **féileacáin** | butterfly | $f^j$ | GEN | SG | INDEF | No |
| **fhéileacáin** | butterfly | Ø | GEN | SG | DEF | Lenition |
| **féileacán** | butterfly | $f^j$ | GEN | PL | INDEF | No |
| **bhféileacán** | butterfly | $v^j$ | GEN | PL | DEF | Eclipsis |
| **fhéileacán** | butterfly | Ø | DAT | SG | INDEF | Lenition |
| **bhféileacán** | butterfly | $v^j$ | DAT | SG | DEF | Eclipsis |
| **fhéileacáin** | butterfly | Ø | DAT | PL | INDEF | Lenition |
| **féileacáin** | butterfly | $f^j$ | DAT | PL | DEF | No |
| **gamhain** | calf | g | NOM | SG | INDEF | No |
| **gamhain** | calf | g | NOM | SG | DEF | No |
| **gamhna** | calf | g | NOM | PL | INDEF | No |
| **gamhna** | calf | g | NOM | PL | DEF | No |
| **gamhna** | calf | g | GEN | SG | INDEF | No |
| **ghamhna** | calf | ɣ | GEN | SG | DEF | Lenition |
| **gamhna** | calf | g | GEN | PL | INDEF | No |
| **ngamhna** | calf | ŋ | GEN | PL | DEF | Eclipsis |
| **ghamhain** | calf | ɣ | DAT | SG | INDEF | Lenition |
| **ngamhain** | calf | ŋ | DAT | SG | DEF | Eclipsis |
| **ghamhna** | calf | ɣ | DAT | PL | INDEF | Lenition |
| **gamhna** | calf | g | DAT | PL | DEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **cat** | cat | k | NOM | SG | INDEF | No |
| **cat** | cat | k | NOM | SG | DEF | No |
| **cait** | cat | k | NOM | PL | INDEF | No |
| **cait** | cat | k | NOM | PL | DEF | No |
| **cait** | cat | k | GEN | SG | INDEF | No |
| **chait** | cat | x | GEN | SG | DEF | Lenition |
| **cat** | cat | k | GEN | PL | INDEF | No |
| **gcat** | cat | g | GEN | PL | DEF | Eclipsis |
| **chat** | cat | x | DAT | SG | INDEF | Lenition |
| **gcat** | cat | g | DAT | SG | DEF | Eclipsis |
| **chait** | cat | x | DAT | PL | INDEF | Lenition |
| **cait** | cat | k | DAT | PL | DEF | No |
| **deán** | channel in strand | $d^j$ | NOM | SG | INDEF | No |
| **deán** | channel in strand | $d^j$ | NOM | SG | DEF | No |
| **deánta** | channel in strand | $d^j$ | NOM | PL | INDEF | No |
| **deánta** | channel in strand | $d^j$ | NOM | PL | DEF | No |
| **deáin** | channel in strand | $d^j$ | GEN | SG | INDEF | No |
| **deáin** | channel in strand | $d^j$ | GEN | SG | DEF | No |
| **deánta** | channel in strand | $d^j$ | GEN | PL | INDEF | No |
| **ndeánta** | channel in strand | $n^j$ | GEN | PL | DEF | Eclipsis |
| **dheán** | channel in strand | $ɣ^j$ | DAT | SG | INDEF | Lenition |
| **deán** | channel in strand | $d^j$ | DAT | SG | DEF | No |
| **dheánta** | channel in strand | $ɣ^j$ | DAT | PL | INDEF | Lenition |
| **deánta** | channel in strand | $d^j$ | DAT | PL | DEF | No |
| **páiste** | child | p | NOM | SG | INDEF | No |
| **páiste** | child | p | NOM | SG | DEF | No |
| **páistí** | child | p | NOM | PL | INDEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **páistí** | child | p | NOM | PL | DEF | No |
| **páiste** | child | p | GEN | SG | INDEF | No |
| **pháiste** | child | f | GEN | SG | DEF | Lenition |
| **páistí** | child | p | GEN | PL | INDEF | No |
| **bpáistí** | child | b | GEN | PL | DEF | Eclipsis |
| **pháiste** | child | f | DAT | SG | INDEF | Lenition |
| **bpáiste** | child | b | DAT | SG | DEF | Eclipsis |
| **pháistí** | child | f | DAT | PL | INDEF | Lenition |
| **páistí** | child | p | DAT | PL | DEF | No |
| **gearán** | complaint | g$^j$ | NOM | SG | INDEF | No |
| **gearán** | complaint | g$^j$ | NOM | SG | DEF | No |
| **gearáin** | complaint | g$^j$ | NOM | PL | INDEF | No |
| **gearáin** | complaint | g$^j$ | NOM | PL | DEF | No |
| **gearáin** | complaint | g$^j$ | GEN | SG | INDEF | No |
| **ghearáin** | complaint | ɣ$^j$ | GEN | SG | DEF | Lenition |
| **gearán** | complaint | g$^j$ | GEN | PL | INDEF | No |
| **ngearán** | complaint | ŋ$^j$ | GEN | PL | DEF | Eclipsis |
| **ghearán** | complaint | ɣ$^j$ | DAT | SG | INDEF | Lenition |
| **ngearán** | complaint | ŋ$^j$ | DAT | SG | DEF | Eclipsis |
| **ghearáin** | complaint | ɣ$^j$ | DAT | PL | INDEF | Lenition |
| **gearáin** | complaint | g$^j$ | DAT | PL | DEF | No |
| **tír** | country | t$^j$ | NOM | SG | INDEF | No |
| **tír** | country | t$^j$ | NOM | SG | DEF | No |
| **tíortha** | country | t$^j$ | NOM | PL | INDEF | No |
| **tíortha** | country | t$^j$ | NOM | PL | DEF | No |
| **tíre** | country | t$^j$ | GEN | SG | INDEF | No |
| **tíre** | country | t$^j$ | GEN | SG | DEF | No |
| **tíortha** | country | t$^j$ | GEN | PL | INDEF | No |
| **dtíortha** | country | d$^j$ | GEN | PL | DEF | Eclipsis |
| **thír** | country | h | DAT | SG | INDEF | Lenition |
| **tír** | country | t$^j$ | DAT | SG | DEF | No |
| **thíortha** | country | h | DAT | PL | INDEF | Lenition |
| **tíortha** | country | t$^j$ | DAT | PL | DEF | No |
| **bó** | cow | b | NOM | SG | INDEF | No |
| **bhó** | cow | v | NOM | SG | DEF | Lenition |
| **ba** | cow | b | NOM | PL | INDEF | No |
| **ba** | cow | b | NOM | PL | DEF | No |
| **bó** | cow | b | GEN | SG | INDEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **bó** | cow | b | GEN | SG | DEF | No |
| **bó** | cow | b | GEN | PL | INDEF | No |
| **mbó** | cow | m | GEN | PL | DEF | Eclipsis |
| **bhó** | cow | v | DAT | SG | INDEF | Lenition |
| **mbó** | cow | m | DAT | SG | DEF | Eclipsis |
| **bha** | cow | v | DAT | PL | INDEF | Lenition |
| **ba** | cow | b | DAT | PL | DEF | No |
| **meabhlaire** | deceiver | $m^j$ | NOM | SG | INDEF | No |
| **meabhlaire** | deceiver | $m^j$ | NOM | SG | DEF | No |
| **meabhlairí** | deceiver | $m^j$ | NOM | PL | INDEF | No |
| **meabhlairí** | deceiver | $m^j$ | NOM | PL | DEF | No |
| **meabhlaire** | deceiver | $m^j$ | GEN | SG | INDEF | No |
| **mheabhlaire** | deceiver | $v^j$ | GEN | SG | DEF | Lenition |
| **meabhlairí** | deceiver | $m^j$ | GEN | PL | INDEF | No |
| **meabhlairí** | deceiver | $m^j$ | GEN | PL | DEF | No |
| **mheabhlaire** | deceiver | $v^j$ | DAT | SG | INDEF | Lenition |
| **meabhlaire** | deceiver | $m^j$ | DAT | SG | DEF | No |
| **mheabhlairí** | deceiver | $v^j$ | DAT | PL | INDEF | Lenition |
| **meabhlairí** | deceiver | $m^j$ | DAT | PL | DEF | No |
| **fia** | deer | $f^j$ | NOM | SG | INDEF | No |
| **fia** | deer | $f^j$ | NOM | SG | DEF | No |
| **fianna** | deer | $f^j$ | NOM | PL | INDEF | No |
| **fianna** | deer | $f^j$ | NOM | PL | DEF | No |
| **fia** | deer | $f^j$ | GEN | SG | INDEF | No |
| **fhia** | deer | Ø | GEN | SG | DEF | Lenition |
| **fianna** | deer | $f^j$ | GEN | PL | INDEF | No |
| **bhfianna** | deer | $v^j$ | GEN | PL | DEF | Eclipsis |
| **fhia** | deer | Ø | DAT | SG | INDEF | Lenition |
| **bhfia** | deer | $v^j$ | DAT | SG | DEF | Eclipsis |
| **fhianna** | deer | Ø | DAT | PL | INDEF | Lenition |
| **fianna** | deer | $f^j$ | DAT | PL | DEF | No |
| **madra** | dog | m | NOM | SG | INDEF | No |
| **madra** | dog | m | NOM | SG | DEF | No |
| **madraí** | dog | m | NOM | PL | INDEF | No |
| **madraí** | dog | m | NOM | PL | DEF | No |
| **madra** | dog | m | GEN | SG | INDEF | No |
| **mhadra** | dog | v | GEN | SG | DEF | Lenition |
| **madraí** | dog | m | GEN | PL | INDEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **madraí** | dog | m | GEN | PL | DEF | No |
| **mhadra** | dog | v | DAT | SG | INDEF | Lenition |
| **madra** | dog | m | DAT | SG | DEF | No |
| **mhadraí** | dog | v | DAT | PL | INDEF | Lenition |
| **madraí** | dog | m | DAT | PL | DEF | No |
| **gúna** | dress | g | NOM | SG | INDEF | No |
| **gúna** | dress | g | NOM | SG | DEF | No |
| **gúnaí** | dress | g | NOM | PL | INDEF | No |
| **gúnaí** | dress | g | NOM | PL | DEF | No |
| **gúna** | dress | g | GEN | SG | INDEF | No |
| **ghúna** | dress | ɣ | GEN | SG | DEF | Lenition |
| **gúnaí** | dress | g | GEN | PL | INDEF | No |
| **ngúnaí** | dress | ŋ | GEN | PL | DEF | Eclipsis |
| **ghúna** | dress | ɣ | DAT | SG | INDEF | Lenition |
| **ngúna** | dress | ŋ | DAT | SG | DEF | Eclipsis |
| **ghúnaí** | dress | ɣ | DAT | PL | INDEF | Lenition |
| **gúnaí** | dress | g | DAT | PL | DEF | No |
| **súil** | eye | s | NOM | SG | INDEF | No |
| **tsúil** | eye | t | NOM | SG | DEF | Lenition |
| **súile** | eye | s | NOM | PL | INDEF | No |
| **súile** | eye | s | NOM | PL | DEF | No |
| **súile** | eye | s | GEN | SG | INDEF | No |
| **súile** | eye | s | GEN | SG | DEF | No |
| **súl** | eye | s | GEN | PL | INDEF | No |
| **súl** | eye | s | GEN | PL | DEF | No |
| **shúil** | eye | h | DAT | SG | INDEF | Lenition |
| **tsúil** | eye | t | DAT | SG | DEF | Lenition |
| **shúile** | eye | h | DAT | PL | INDEF | Lenition |
| **súile** | eye | s | DAT | PL | DEF | No |
| **ceo** | fog | $k^j$ | NOM | SG | INDEF | No |
| **ceo** | fog | $k^j$ | NOM | SG | DEF | No |
| **ceonna** | fog | $k^j$ | NOM | PL | INDEF | No |
| **ceonna** | fog | $k^j$ | NOM | PL | DEF | No |
| **ceo** | fog | $k^j$ | GEN | SG | INDEF | No |
| **cheo** | fog | $x^j$ | GEN | SG | DEF | Lenition |
| **ceonna** | fog | $k^j$ | GEN | PL | INDEF | No |
| **gceonna** | fog | $g^j$ | GEN | PL | DEF | Eclipsis |
| **cheo** | fog | $x^j$ | DAT | SG | INDEF | Lenition |

| | | | | | | |
|---|---|---|---|---|---|---|
| **gceo** | fog | $g^j$ | DAT | SG | DEF | Eclipsis |
| **cheonna** | fog | $x^j$ | DAT | PL | INDEF | Lenition |
| **ceonna** | fog | $k^j$ | DAT | PL | DEF | No |
| **sionnach** | fox | $s^j$ | NOM | SG | INDEF | No |
| **sionnach** | fox | $s^j$ | NOM | SG | DEF | No |
| **sionnaigh** | fox | $s^j$ | NOM | PL | INDEF | No |
| **sionnaigh** | fox | $s^j$ | NOM | PL | DEF | No |
| **sionnaigh** | fox | $s^j$ | GEN | SG | INDEF | No |
| **tsionnaigh** | fox | $t^j$ | GEN | SG | DEF | Lenition |
| **sionnach** | fox | $s^j$ | GEN | PL | INDEF | No |
| **sionnach** | fox | $s^j$ | GEN | PL | DEF | No |
| **shionnach** | fox | h | DAT | SG | INDEF | Lenition |
| **sionnach** | fox | $s^j$ | DAT | SG | DEF | No |
| **shionnaigh** | fox | h | DAT | PL | INDEF | Lenition |
| **sionnaigh** | fox | $s^j$ | DAT | PL | DEF | No |
| **cara** | friend | k | NOM | SG | INDEF | No |
| **cara** | friend | k | NOM | SG | DEF | No |
| **cairde** | friend | k | NOM | PL | INDEF | No |
| **cairde** | friend | k | NOM | PL | DEF | No |
| **carad** | friend | k | GEN | SG | INDEF | No |
| **charad** | friend | x | GEN | SG | DEF | Lenition |
| **cairde** | friend | k | GEN | PL | INDEF | No |
| **gcairde** | friend | g | GEN | PL | DEF | Eclipsis |
| **chara** | friend | x | DAT | SG | INDEF | Lenition |
| **gcara** | friend | g | DAT | SG | DEF | Eclipsis |
| **chairde** | friend | x | DAT | PL | INDEF | Lenition |
| **cairde** | friend | k | DAT | PL | DEF | No |
| **toradh** | fruit | t | NOM | SG | INDEF | No |
| **toradh** | fruit | t | NOM | SG | DEF | No |
| **torthaí** | fruit | t | NOM | PL | INDEF | No |
| **torthaí** | fruit | t | NOM | PL | DEF | No |
| **toraidh** | fruit | t | GEN | SG | INDEF | No |
| **toraidh** | fruit | t | GEN | SG | DEF | No |
| **torthaí** | fruit | t | GEN | PL | INDEF | No |
| **dtorthaí** | fruit | d | GEN | PL | DEF | Eclipsis |
| **thoradh** | fruit | h | DAT | SG | INDEF | Lenition |
| **toradh** | fruit | t | DAT | SG | DEF | No |
| **thorthaí** | fruit | h | DAT | PL | INDEF | Lenition |

| | | | | | | |
|---|---|---|---|---|---|---|
| **torthaí** | fruit | t | DAT | PL | DEF | No |
| **gineadóir** | generator | $g^j$ | NOM | SG | INDEF | No |
| **gineadóir** | generator | $g^j$ | NOM | SG | DEF | No |
| **gineadóirí** | generator | $g^j$ | NOM | PL | INDEF | No |
| **gineadóirí** | generator | $g^j$ | NOM | PL | DEF | No |
| **gineadóra** | generator | $g^j$ | GEN | SG | INDEF | No |
| **ghineadóra** | generator | ɣ$^j$ | GEN | SG | DEF | Lenition |
| **gineadóirí** | generator | $g^j$ | GEN | PL | INDEF | No |
| **ngineadóirí** | generator | ŋ$^j$ | GEN | PL | DEF | Eclipsis |
| **ghineadóir** | generator | ɣ$^j$ | DAT | SG | INDEF | Lenition |
| **ngineadóir** | generator | ŋ$^j$ | DAT | SG | DEF | Eclipsis |
| **ghineadóirí** | generator | ɣ$^j$ | DAT | PL | INDEF | Lenition |
| **gineadóirí** | generator | $g^j$ | DAT | PL | DEF | No |
| **dia** | god | $d^j$ | NOM | SG | INDEF | No |
| **dia** | god | $d^j$ | NOM | SG | DEF | No |
| **déithe** | god | $d^j$ | NOM | PL | INDEF | No |
| **déithe** | god | $d^j$ | NOM | PL | DEF | No |
| **dé** | god | $d^j$ | GEN | SG | INDEF | No |
| **dé** | god | $d^j$ | GEN | SG | DEF | No |
| **déithe** | god | $d^j$ | GEN | PL | INDEF | No |
| **ndéithe** | god | $n^j$ | GEN | PL | DEF | Eclipsis |
| **dhia** | god | ɣ$^j$ | DAT | SG | INDEF | Lenition |
| **dia** | god | $d^j$ | DAT | SG | DEF | No |
| **dhéithe** | god | ɣ$^j$ | DAT | PL | INDEF | Lenition |
| **déithe** | god | $d^j$ | DAT | PL | DEF | No |
| **lámh** | hand | l | NOM | SG | INDEF | No |
| **lámh** | hand | l | NOM | SG | DEF | No |
| **lámha** | hand | l | NOM | PL | INDEF | No |
| **lámha** | hand | l | NOM | PL | DEF | No |
| **láimhe** | hand | l | GEN | SG | INDEF | No |
| **láimhe** | hand | l | GEN | SG | DEF | No |
| **lámh** | hand | l | GEN | PL | INDEF | No |
| **lámh** | hand | l | GEN | PL | DEF | No |
| **lámh** | hand | l | DAT | SG | INDEF | No |
| **lámh** | hand | l | DAT | SG | DEF | No |
| **lámha** | hand | l | DAT | PL | INDEF | No |
| **lámha** | hand | l | DAT | PL | DEF | No |
| **méadú** | increase | $m^j$ | NOM | SG | INDEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **méadú** | increase | $m^j$ | NOM | SG | DEF | No |
| **méaduithe** | increase | $m^j$ | NOM | PL | INDEF | No |
| **méaduithe** | increase | $m^j$ | NOM | PL | DEF | No |
| **méadaithe** | increase | $m^j$ | GEN | SG | INDEF | No |
| **mhéadaithe** | increase | $v^j$ | GEN | SG | DEF | Lenition |
| **méaduithe** | increase | $m^j$ | GEN | PL | INDEF | No |
| **méaduithe** | increase | $m^j$ | GEN | PL | DEF | No |
| **mhéadú** | increase | $v^j$ | DAT | SG | INDEF | Lenition |
| **méadú** | increase | $m^j$ | DAT | SG | DEF | No |
| **mheáduithe** | increase | $v^j$ | DAT | PL | INDEF | Lenition |
| **méaduithe** | increase | $m^j$ | DAT | PL | DEF | No |
| **post** | job | p | NOM | SG | INDEF | No |
| **post** | job | p | NOM | SG | DEF | No |
| **poist** | job | p | NOM | PL | INDEF | No |
| **poist** | job | p | NOM | PL | DEF | No |
| **poist** | job | p | GEN | SG | INDEF | No |
| **phoist** | job | f | GEN | SG | DEF | Lenition |
| **post** | job | p | GEN | PL | INDEF | No |
| **bpost** | job | b | GEN | PL | DEF | Eclipsis |
| **phost** | job | f | DAT | SG | INDEF | Lenition |
| **bpost** | job | b | DAT | SG | DEF | Eclipsis |
| **phoist** | job | f | DAT | PL | INDEF | Lenition |
| **poist** | job | p | DAT | PL | DEF | No |
| **fadhb** | knot | f | NOM | SG | INDEF | No |
| **fhadhb** | knot | Ø | NOM | SG | DEF | Lenition |
| **fadhbanna** | knot | f | NOM | PL | INDEF | No |
| **fadhbanna** | knot | f | NOM | PL | DEF | No |
| **faidhbe** | knot | f | GEN | SG | INDEF | No |
| **faidhbe** | knot | f | GEN | SG | DEF | No |
| **fadhbanna** | knot | f | GEN | PL | INDEF | No |
| **bhfadhbanna** | knot | v | GEN | PL | DEF | Eclipsis |
| **fhadhb** | knot | Ø | DAT | SG | INDEF | Lenition |
| **bhfadhb** | knot | v | DAT | SG | DEF | Eclipsis |
| **fhadhbanna** | knot | Ø | DAT | PL | INDEF | Lenition |
| **fadhbanna** | knot | f | DAT | PL | DEF | No |
| **teanga** | language | $t^j$ | NOM | SG | INDEF | No |
| **teanga** | language | $t^j$ | NOM | SG | DEF | No |
| **teangacha** | language | $t^j$ | NOM | PL | INDEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **teangacha** | language | $t^j$ | NOM | PL | DEF | No |
| **teanga** | language | $t^j$ | GEN | SG | INDEF | No |
| **teanga** | language | $t^j$ | GEN | SG | DEF | No |
| **teangacha** | language | $t^j$ | GEN | PL | INDEF | No |
| **dteangacha** | language | $d^j$ | GEN | PL | DEF | Eclipsis |
| **theanga** | language | h | DAT | SG | INDEF | Lenition |
| **teanga** | language | $t^j$ | DAT | SG | DEF | No |
| **theangacha** | language | h | DAT | PL | INDEF | Lenition |
| **teangacha** | language | $t^j$ | DAT | PL | DEF | No |
| **ceadúnas** | license | $k^j$ | NOM | SG | INDEF | No |
| **ceadúnas** | license | $k^j$ | NOM | SG | DEF | No |
| **ceadúnais** | license | $k^j$ | NOM | PL | INDEF | No |
| **ceadúnais** | license | $k^j$ | NOM | PL | DEF | No |
| **ceadúnais** | license | $k^j$ | GEN | SG | INDEF | No |
| **cheadúnais** | license | $x^j$ | GEN | SG | DEF | Lenition |
| **ceadúnas** | license | $k^j$ | GEN | PL | INDEF | No |
| **gceadúnas** | license | $g^j$ | GEN | PL | DEF | Eclipsis |
| **cheadúnas** | license | $x^j$ | DAT | SG | INDEF | Lenition |
| **gceadúnas** | license | $g^j$ | DAT | SG | DEF | Eclipsis |
| **cheadúnais** | license | $x^j$ | DAT | PL | INDEF | Lenition |
| **ceadúnais** | license | $k^j$ | DAT | PL | DEF | No |
| **lictéar** | lighter | $l^j$ | NOM | SG | INDEF | No |
| **lictéar** | lighter | $l^j$ | NOM | SG | DEF | No |
| **lictéir** | lighter | $l^j$ | NOM | PL | INDEF | No |
| **lictéir** | lighter | $l^j$ | NOM | PL | DEF | No |
| **lictéir** | lighter | $l^j$ | GEN | SG | INDEF | No |
| **lictéir** | lighter | $l^j$ | GEN | SG | DEF | No |
| **lictéar** | lighter | $l^j$ | GEN | PL | INDEF | No |
| **lictéar** | lighter | $l^j$ | GEN | PL | DEF | No |
| **lictéar** | lighter | $l^j$ | DAT | SG | INDEF | No |
| **lictéar** | lighter | $l^j$ | DAT | SG | DEF | No |
| **lictéir** | lighter | $l^j$ | DAT | PL | INDEF | No |
| **lictéir** | lighter | $l^j$ | DAT | PL | DEF | No |
| **Bealtaine** | May (month) | $b^j$ | NOM | SG | INDEF | No |
| **Bhealtaine** | May (month) | $v^j$ | NOM | SG | DEF | Lenition |
| **Bealtainí** | May (month) | $b^j$ | NOM | PL | INDEF | No |
| **Bealtainí** | May (month) | $b^j$ | NOM | PL | DEF | No |
| **Bealtaine** | May (month) | $b^j$ | GEN | SG | INDEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Bealtaine** | May (month) | $b^j$ | GEN | SG | DEF | No |
| **Bealtainí** | May (month) | $b^j$ | GEN | PL | INDEF | No |
| **mBealtainí** | May (month) | $m^j$ | GEN | PL | DEF | Eclipsis |
| **Bhealtaine** | May (month) | $v^j$ | DAT | SG | INDEF | Lenition |
| **mBealtaine** | May (month) | $m^j$ | DAT | SG | DEF | Eclipsis |
| **Bhealtainí** | May (month) | $v^j$ | DAT | PL | INDEF | Lenition |
| **Bealtainí** | May (month) | $b^j$ | DAT | PL | DEF | No |
| **luch** | mouse | l | NOM | SG | INDEF | No |
| **luch** | mouse | l | NOM | SG | DEF | No |
| **lucha** | mouse | l | NOM | PL | INDEF | No |
| **lucha** | mouse | l | NOM | PL | DEF | No |
| **luiche** | mouse | l | GEN | SG | INDEF | No |
| **luiche** | mouse | l | GEN | SG | DEF | No |
| **luch** | mouse | l | GEN | PL | INDEF | No |
| **luch** | mouse | l | GEN | PL | DEF | No |
| **luch** | mouse | l | DAT | SG | INDEF | No |
| **luch** | mouse | l | DAT | SG | DEF | No |
| **lucha** | mouse | l | DAT | PL | INDEF | No |
| **lucha** | mouse | l | DAT | PL | DEF | No |
| **néarchóras** | nervous system | $n^j$ | NOM | SG | INDEF | No |
| **néarchóras** | nervous system | $n^j$ | NOM | SG | DEF | No |
| **néarchórais** | nervous system | $n^j$ | NOM | PL | INDEF | No |
| **néarchórais** | nervous system | $n^j$ | NOM | PL | DEF | No |
| **néarchórais** | nervous system | $n^j$ | GEN | SG | INDEF | No |
| **néarchórais** | nervous system | $n^j$ | GEN | SG | DEF | No |
| **néarchóras** | nervous system | $n^j$ | GEN | PL | INDEF | No |
| **néarchóras** | nervous system | $n^j$ | GEN | PL | DEF | No |
| **néarchóras** | nervous system | $n^j$ | DAT | SG | INDEF | No |
| **néarchóras** | nervous system | $n^j$ | DAT | SG | DEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **néarchórais** | nervous system | $n^j$ | DAT | PL | INDEF | No |
| **néarchórais** | nervous system | $n^j$ | DAT | PL | DEF | No |
| **nuachtán** | newspaper | n | NOM | SG | INDEF | No |
| **nuachtán** | newspaper | n | NOM | SG | DEF | No |
| **nuachtáin** | newspaper | n | NOM | PL | INDEF | No |
| **nuachtáin** | newspaper | n | NOM | PL | DEF | No |
| **nuachtáin** | newspaper | n | GEN | SG | INDEF | No |
| **nuachtáin** | newspaper | n | GEN | SG | DEF | No |
| **nuachtán** | newspaper | n | GEN | PL | INDEF | No |
| **nuachtán** | newspaper | n | GEN | PL | DEF | No |
| **nuachtán** | newspaper | n | DAT | SG | INDEF | No |
| **nuachtán** | newspaper | n | DAT | SG | DEF | No |
| **nuachtáin** | newspaper | n | DAT | PL | INDEF | No |
| **nuachtáin** | newspaper | n | DAT | PL | DEF | No |
| **Samhain** | November | s | NOM | SG | INDEF | No |
| **tSamhain** | November | t | NOM | SG | DEF | Lenition |
| **Samhnacha** | November | s | NOM | PL | INDEF | No |
| **Samhnacha** | November | s | NOM | PL | DEF | No |
| **Samhna** | November | s | GEN | SG | INDEF | No |
| **Samhna** | November | s | GEN | SG | DEF | No |
| **Samhnacha** | November | s | GEN | PL | INDEF | No |
| **Samhnacha** | November | s | GEN | PL | DEF | No |
| **Shamhain** | November | h | DAT | SG | INDEF | Lenition |
| **tSamhain** | November | t | DAT | SG | DEF | Lenition |
| **Shamhnacha** | November | h | DAT | PL | INDEF | Lenition |
| **Samhnacha** | November | s | DAT | PL | DEF | No |
| **piobar** | pepper | $p^j$ | NOM | SG | INDEF | No |
| **piobar** | pepper | $p^j$ | NOM | SG | DEF | No |
| **piobair** | pepper | $p^j$ | NOM | PL | INDEF | No |
| **piobair** | pepper | $p^j$ | NOM | PL | DEF | No |
| **piobair** | pepper | $p^j$ | GEN | SG | INDEF | No |
| **phiobair** | pepper | $f^j$ | GEN | SG | DEF | Lenition |
| **piobar** | pepper | $p^j$ | GEN | PL | INDEF | No |
| **bpiobar** | pepper | $b^j$ | GEN | PL | DEF | Eclipsis |
| **phiobar** | pepper | $f^j$ | DAT | SG | INDEF | Lenition |
| **bpiobar** | pepper | $b^j$ | DAT | SG | DEF | Eclipsis |

| | | | | | | |
|---|---|---|---|---|---|---|
| **phiobair** | pepper | $f^j$ | DAT | PL | INDEF | Lenition |
| **piobair** | pepper | $p^j$ | DAT | PL | DEF | No |
| **peiteal** | petal | $p^j$ | NOM | SG | INDEF | No |
| **peiteal** | petal | $p^j$ | NOM | SG | DEF | No |
| **peitil** | petal | $p^j$ | NOM | PL | INDEF | No |
| **peitil** | petal | $p^j$ | NOM | PL | DEF | No |
| **peitil** | petal | $p^j$ | GEN | SG | INDEF | No |
| **pheitil** | petal | $f^j$ | GEN | SG | DEF | Lenition |
| **peiteal** | petal | $p^j$ | GEN | PL | INDEF | No |
| **bpeiteal** | petal | $b^j$ | GEN | PL | DEF | Eclipsis |
| **pheiteal** | petal | $f^j$ | DAT | SG | INDEF | Lenition |
| **bpeiteal** | petal | $b^j$ | DAT | SG | DEF | Eclipsis |
| **pheitil** | petal | $f^j$ | DAT | PL | INDEF | Lenition |
| **peitil** | petal | $p^j$ | DAT | PL | DEF | No |
| **dán** | poem | d | NOM | SG | INDEF | No |
| **dán** | poem | d | NOM | SG | DEF | No |
| **dánta** | poem | d | NOM | PL | INDEF | No |
| **dánta** | poem | d | NOM | PL | DEF | No |
| **dáin** | poem | d | GEN | SG | INDEF | No |
| **dáin** | poem | d | GEN | SG | DEF | No |
| **dánta** | poem | d | GEN | PL | INDEF | No |
| **ndánta** | poem | n | GEN | PL | DEF | Eclipsis |
| **dhán** | poem | ɣ | DAT | SG | INDEF | Lenition |
| **dán** | poem | d | DAT | SG | DEF | No |
| **dhánta** | poem | ɣ | DAT | PL | INDEF | Lenition |
| **dánta** | poem | d | DAT | PL | DEF | No |
| **rón** | seal | r | NOM | SG | INDEF | No |
| **rón** | seal | r | NOM | SG | DEF | No |
| **rónta** | seal | r | NOM | PL | INDEF | No |
| **rónta** | seal | r | NOM | PL | DEF | No |
| **róin** | seal | r | GEN | SG | INDEF | No |
| **róin** | seal | r | GEN | SG | DEF | No |
| **rónta** | seal | r | GEN | PL | INDEF | No |
| **rónta** | seal | r | GEN | PL | DEF | No |
| **rón** | seal | r | DAT | SG | INDEF | No |
| **rón** | seal | r | DAT | SG | DEF | No |
| **rónta** | seal | r | DAT | PL | INDEF | No |
| **rónta** | seal | r | DAT | PL | DEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **siopa** | shop | $s^j$ | NOM | SG | INDEF | No |
| **siopa** | shop | $s^j$ | NOM | SG | DEF | No |
| **siopaí** | shop | $s^j$ | NOM | PL | INDEF | No |
| **siopaí** | shop | $s^j$ | NOM | PL | DEF | No |
| **siopa** | shop | $s^j$ | GEN | SG | INDEF | No |
| **tsiopa** | shop | $t^j$ | GEN | SG | DEF | Lenition |
| **siopaí** | shop | $s^j$ | GEN | PL | INDEF | No |
| **siopaí** | shop | $s^j$ | GEN | PL | DEF | No |
| **shiopa** | shop | h | DAT | SG | INDEF | Lenition |
| **siopa** | shop | $s^j$ | DAT | SG | DEF | No |
| **shiopaí** | shop | h | DAT | PL | INDEF | Lenition |
| **siopaí** | shop | $s^j$ | DAT | PL | DEF | No |
| **ríog** | spasm | $r^j$ | NOM | SG | INDEF | No |
| **ríog** | spasm | $r^j$ | NOM | SG | DEF | No |
| **ríoga** | spasm | $r^j$ | NOM | PL | INDEF | No |
| **ríoga** | spasm | $r^j$ | NOM | PL | DEF | No |
| **ríge** | spasm | $r^j$ | GEN | SG | INDEF | No |
| **ríge** | spasm | $r^j$ | GEN | SG | DEF | No |
| **ríog** | spasm | $r^j$ | GEN | PL | INDEF | No |
| **ríog** | spasm | $r^j$ | GEN | PL | DEF | No |
| **ríog** | spasm | $r^j$ | DAT | SG | INDEF | No |
| **ríog** | spasm | $r^j$ | DAT | SG | DEF | No |
| **ríoga** | spasm | $r^j$ | DAT | PL | INDEF | No |
| **ríoga** | spasm | $r^j$ | DAT | PL | DEF | No |
| **réalta** | star | $r^j$ | NOM | SG | INDEF | No |
| **réalta** | star | $r^j$ | NOM | SG | DEF | No |
| **réaltaí** | star | $r^j$ | NOM | PL | INDEF | No |
| **réaltaí** | star | $r^j$ | NOM | PL | DEF | No |
| **réalta** | star | $r^j$ | GEN | SG | INDEF | No |
| **réalta** | star | $r^j$ | GEN | SG | DEF | No |
| **réaltaí** | star | $r^j$ | GEN | PL | INDEF | No |
| **réaltaí** | star | $r^j$ | GEN | PL | DEF | No |
| **réalta** | star | $r^j$ | DAT | SG | INDEF | No |
| **réalta** | star | $r^j$ | DAT | SG | DEF | No |
| **réaltaí** | star | $r^j$ | DAT | PL | INDEF | No |
| **réaltaí** | star | $r^j$ | DAT | PL | DEF | No |
| **bocóid** | stud/boss | b | NOM | SG | INDEF | No |
| **bhocóid** | stud/boss | v | NOM | SG | DEF | Lenition |

| | | | | | | |
|---|---|---|---|---|---|---|
| **bocóidí** | stud/boss | b | NOM | PL | INDEF | No |
| **bocóidí** | stud/boss | b | NOM | PL | DEF | No |
| **bocóide** | stud/boss | b | GEN | SG | INDEF | No |
| **bocóide** | stud/boss | b | GEN | SG | DEF | No |
| **bocóidí** | stud/boss | b | GEN | PL | INDEF | No |
| **mbocóidí** | stud/boss | m | GEN | PL | DEF | Eclipsis |
| **bhocóid** | stud/boss | v | DAT | SG | INDEF | Lenition |
| **mbocóid** | stud/boss | m | DAT | SG | DEF | Eclipsis |
| **bhocóidí** | stud/boss | v | DAT | PL | INDEF | Lenition |
| **bocóidí** | stud/boss | b | DAT | PL | DEF | No |
| **dalta** | student | d | NOM | SG | INDEF | No |
| **dalta** | student | d | NOM | SG | DEF | No |
| **daltaí** | student | d | NOM | PL | INDEF | No |
| **daltaí** | student | d | NOM | PL | DEF | No |
| **dalta** | student | d | GEN | SG | INDEF | No |
| **dalta** | student | d | GEN | SG | DEF | No |
| **daltaí** | student | d | GEN | PL | INDEF | No |
| **ndaltaí** | student | n | GEN | PL | DEF | Eclipsis |
| **dhalta** | student | ɣ | DAT | SG | INDEF | Lenition |
| **dalta** | student | d | DAT | SG | DEF | No |
| **dhaltaí** | student | ɣ | DAT | PL | INDEF | Lenition |
| **daltaí** | student | d | DAT | PL | DEF | No |
| **tábhairne** | tavern | t | NOM | SG | INDEF | No |
| **tábhairne** | tavern | t | NOM | SG | DEF | No |
| **tábhairní** | tavern | t | NOM | PL | INDEF | No |
| **tábhairní** | tavern | t | NOM | PL | DEF | No |
| **tábhairne** | tavern | t | GEN | SG | INDEF | No |
| **tábhairne** | tavern | t | GEN | SG | DEF | No |
| **tábhairní** | tavern | t | GEN | PL | INDEF | No |
| **dtábhairní** | tavern | d | GEN | PL | DEF | Eclipsis |
| **thábhairne** | tavern | h | DAT | SG | INDEF | Lenition |
| **tábhairne** | tavern | t | DAT | SG | DEF | No |
| **thábhairní** | tavern | h | DAT | PL | INDEF | Lenition |
| **tábhairní** | tavern | t | DAT | PL | DEF | No |
| **nós** | tradition | n | NOM | SG | INDEF | No |
| **nós** | tradition | n | NOM | SG | DEF | No |
| **nósanna** | tradition | n | NOM | PL | INDEF | No |
| **nósanna** | tradition | n | NOM | PL | DEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **nóis** | tradition | n | GEN | SG | INDEF | No |
| **nóis** | tradition | n | GEN | SG | DEF | No |
| **nósanna** | tradition | n | GEN | PL | INDEF | No |
| **nósanna** | tradition | n | GEN | PL | DEF | No |
| **nós** | tradition | n | DAT | SG | INDEF | No |
| **nós** | tradition | n | DAT | SG | DEF | No |
| **nósanna** | tradition | n | DAT | PL | INDEF | No |
| **nósanna** | tradition | n | DAT | PL | DEF | No |
| **niteoir** | washer (person) | $n^j$ | NOM | SG | INDEF | No |
| **niteoir** | washer (person) | $n^j$ | NOM | SG | DEF | No |
| **niteoirí** | washer (person) | $n^j$ | NOM | PL | INDEF | No |
| **niteoirí** | washer (person) | $n^j$ | NOM | PL | DEF | No |
| **niteora** | washer (person) | $n^j$ | GEN | SG | INDEF | No |
| **niteora** | washer (person) | $n^j$ | GEN | SG | DEF | No |
| **niteoirí** | washer (person) | $n^j$ | GEN | PL | INDEF | No |
| **niteoirí** | washer (person) | $n^j$ | GEN | PL | DEF | No |
| **niteoir** | washer (person) | $n^j$ | DAT | SG | INDEF | No |
| **niteoir** | washer (person) | $n^j$ | DAT | SG | DEF | No |
| **niteoirí** | washer (person) | $n^j$ | DAT | PL | INDEF | No |
| **niteoirí** | washer (person) | $n^j$ | DAT | PL | DEF | No |
| **fuinneog** | window | f | NOM | SG | INDEF | No |
| **fhuinneog** | window | Ø | NOM | SG | DEF | Lenition |
| **fuinneoga** | window | f | NOM | PL | INDEF | No |
| **fuinneoga** | window | f | NOM | PL | DEF | No |
| **fuinneoige** | window | f | GEN | SG | INDEF | No |
| **fuinneoige** | window | f | GEN | SG | DEF | No |
| **fuinneog** | window | f | GEN | PL | INDEF | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| **bhfuinneog** | window | v | GEN | PL | DEF | Eclipsis |
| **fhuinneog** | window | Ø | DAT | SG | INDEF | Lenition |
| **bhfuinneog** | window | v | DAT | SG | DEF | Eclipsis |
| **fhuinneoga** | window | Ø | DAT | PL | INDEF | Lenition |
| **fuinneoga** | window | f | DAT | PL | DEF | No |