Assessed and Annotated Vowel Lengths in Spoken Icelandic Sentences for L1 and L2 Speakers: A Resource for Pronunciation Training

Caitlin Laura Richter
Reykjavik University
caitlinr@ru.is

Kolbrún Friðriksdóttir
University of Iceland
kolbrunf@hi.is

Kormákur Logi Bergsson
University of Iceland
klb16@hi.is

Erik Anders Maher
University of Iceland
eam16@hi.is

Ragnheiður María Benediktsdóttir
University of Iceland
rmb9@hi.is

Jon Gudnason
Reykjavik University
jg@ru.is

Abstract

We introduce a dataset of time-aligned phonetic transcriptions focusing on vowel length (quantity) in Icelandic. Ultimately, this aims to support computer assisted pronunciation training (CAPT) software, to automatically assess length and possible errors in Icelandic learners' pronunciations. The dataset contains a range of long and short vowel targets, including the first acoustic description of quantity in non-native Icelandic. Evaluations assess how manual annotations and automatic forced alignment characterise quantity contrasts. Initial analyses also imply partial acquisition of phonologically conditioned quantity alternations by non-native speakers.

1 Introduction

We present a corpus of Icelandic speech with manually corrected time-aligned phonetic transcriptions, targeted towards native and non-native Icelandic speakers' acoustic realisations of vowel quantity (length). Quantity is important in non-native (L2) Icelandic learning because it is contrastive, as in *vinur* [vɪːnʏr] 'friend', *vinnur* [vɪnːʏr] 'you, s/he work(s)', but challenging for many learners whose first languages do not use this cue. Computer assisted language learning (CALL) such as pronunciation training (CAPT) enables self-directed learning beyond traditional classrooms, and could provide opportunities to practice and internalise the Icelandic quantity system.

The acoustic implementation of Icelandic quantity has been studied only in small manually annotated native-speaker (L1) datasets. Addressing learners' needs requires (i) understanding quantity realisation in a broad sample of L1 and L2 speech, and (ii) developing scalable automated methods to describe a sufficient sample of the language and to evaluate learners' speech relative to acoustic targets in autonomous interactive CAPT software.

We release time-aligned phonetic annotations for 2707 tokens of 72 Icelandic words,[1] greatly increasing the variety of contexts with available acoustic data on quantity, and including non-native speech for the first time. §4 uses this data to explore the realisation of quantity contrasts, comparing manual annotations and automated equivalents from the Montreal Forced Aligner (MFA), to address four **Research questions**:

RQ1 How do (subsets of) the annotated data relate to expectations from comparable studies?
RQ2 How strongly do quantity contrasts emerge in the annotated features, for L1 and L2 speakers?
RQ3 How accurate is Montreal Forced Aligner (MFA) timing, compared to gold annotations?
RQ4 How useful is MFA for issues in RQs 1-2?

2 Vowel Quantity in Icelandic

2.1 Language description

Stressed vowels in Icelandic, generally the first syllable of a word, have a quantity contrast conditioned by the vowel's environment (Einarsson, 1945; Kristinsson et al., 1985). A usual description of surface facts (Árnason, 1998; Gussmann, 2011) is that stressed vowels (including diphthongs) are long when followed by at most one consonant: *tré* 'tree', *hús* 'house', the first vowel *í* in *sími* 'telephone'. They are short when two or more consonants (geminates included) follow them before either the next vowel or the end of the word, e.g. *mjólk* 'milk', *a* in *pabbi* 'dad', except that specific clusters {p,t,k,s}+{j,v,r} are preceded by long vowels, e.g. long *i* in *sitja* 'sit'. In phonological terms it is conventional to say that vowels are long in open syllables and closed in short syllables, but it has proved challenging to complete this with an account of Icelandic syllable structure that does not

[1] https://github.com/catiR/length-contrast-data-isl

Proceedings of the Joint 25th Nordic Conference on Computational Linguistics and 11th Baltic Conference on Human Language Technologies (NoDaLiDa/Baltic-HLT 2025), pages 518–524

circularly refer back to vowel length (see Árnason 2011; Craioveanu 2023; Fortuna 2016; Gussmann 2011; Þráinsson 1994; for issues bearing on phonological characterisation and the interface with morphosyntax). In practise, language teachers as well as linguists presenting the most thorough descriptions of Icelandic vowel length rarely complete formal phonological accounts of it (Árnason, 1998; Kristinsson, 1988; Craioveanu, 2023), so we continue the convenience of using the orthography as the simplest means to communicate.

2.2 Acoustic properties

The reader is referred to Pind (1999) for a review of acoustic research on Icelandic vowel quantity from Einarsson (1927) onwards, and subsequently Árnason (2011). In summary, absolute durations of long vs. short vowel segments overlap considerably, but there is a complementary relationship between vowels and the consonant(s) that follow them, such that these segments' combined duration in a word is relatively consistent: [aːl] in *gala* and [alː] in *galla* (Pind, 1995; Einarsson, 1927). Therefore, Icelandic vowel quantity is often described by a proportion, formulated as V/(V+C), the ratio of vowel duration to total vowel+consonant durations (Pind, 1995); this calculation variously incorporates segments from either one or two syllables, as consonants in C are in either the coda of the stressed syllable or the onset of the next. Properties like vowel quality have also been identified as secondary cues to quantity for some vowels (Pind, 1999; Kristinsson et al., 1985). However, the acoustic research draws on narrowly restricted samples of few or one speaker(s), minimal vowel/syllable types, or only sentence-initial words. Audio and annotations are generally not accessible, and much in the language remains undescribed, such as any diphthongs, or L2 speech.

2.3 Teaching vowel quantity

Perceiving and producing quantity contrasts, as in *koma* 'come', *komma* 'comma', can be challenging for students of L2 Icelandic whose native language lacks such contrasts (McAllister et al., 2002). Computer assisted pronunciation training (CAPT) can offer help such as interactive exercises with feedback (Arnbjörnsdóttir et al., 2020; Bédi, 2022). Pronunciation accuracy assessment has been developed in coordination with lesson content of the free course *Icelandic Online*, but this does not give feedback on quantity errors, which is difficult to provide without knowing what learners' acoustic targets are (Bédi, 2022; Bedi et al., 2024).

3 Corpus creation

3.1 Speech data

Audio is drawn from Samrómur, Samrómur Queries, Samrómur Unverified, and Samrómur L2 (Mollberg et al., 2021; Hedström et al., 2021, 2022a,b), recorded from 2019 onwards by native and non-native Icelandic speakers. Excluding child recordings (under age 18) there are in total 1.4 million sentences and 180,598 unique word types in over 1000 hours of speech. As corpora of crowdsourced read sentences, these are typical of audio conditions that pronunciation training software processes for CAPT users.

Icelandic language proficiency levels and native language backgrounds of L2 Icelandic learners in these corpora are not reported, but plenty of variation in both of these factors was subjectively observed during manual annotation. Overall accuracy of phoneme reproduction and reading suggests that many speakers are intermediate to advanced learners of the language, although some speakers are likely within their first year of study and in certain recordings the speaker's prosody implies failure to semantically understand the sentence. Occasional deviations from Icelandic L1 pronunciation shown by L2 speakers were noted in vowel length and quality, with some relation to apparent first language background.

3.2 Target words

72 words of interest were sampled in two rounds of annotation. A complete list is provided in Appendix A.

The initial *validation* sample (36 words) is parallel to Experiment 2 from Pind (1999) and Experiment 1 of Pind (1995). In the former, 25 speakers read target words *saki, saggi, seki, seggi* within a paragraph; the reading context and number of speakers stand out as a clear choice for comparison to Samrómur data. From the latter, data on *kala, gala, Kalla, galla* (Pind, 1995) includes fewer speakers, but has similar enough acoustic analysis to also draw into comparison. Some of these 8 words are very infrequent, so to better assess reliability and variability, the validation sample is filled out with other two-syllable words that differ from Pind's only in the word onset, e.g. *tala, aggi, dreki*.

The second *extension* sample (36 words) highlights variation in stressed vowel phenomena, including: diphthongs; a range of vowels preceding different consonantal contexts such as nasals, fricatives, short and long trill, and assorted clusters; words with 'exceptional' consonant clusters preceded by long vowels; and quantity alternations within a morpheme as conditioned by compounding, inflection, and/or vowel syncope.

For each of the two samples, the most frequent words matching criteria were selected from Samrómur data. Annotators checked and filtered each word's carrier sentences (in case of homonyms with different pronunciations), and where possible annotated at most 10 tokens from the same carrier sentence per L1/L2 speaker group.

3.3 Forced Alignment

As fully manual phonetic transcription is excessively time consuming, data was preprocessed by forced alignment, which annotators reviewed and corrected. The Montreal Forced Aligner (MFA) is a widely used toolkit built on Kaldi with standard GMM-HMM triphone acoustic models (McAuliffe et al., 2017). We train the aligner's acoustic models on 20 hours of Icelandic speech from Samrómur, and use the General Icelandic Pronunciation Dictionary for ASR (Nikulásdóttir and Guðnason, 2017), to which a few target words not already present were manually added.

3.4 Annotation

Recordings were annotated by three of the authors while enrolled in undergraduate degrees on linguistics and/or Icelandic language at the University of Iceland. Two annotators are native Icelandic speakers and all have training in Icelandic phonetics. Annotation was carried out by reviewing and adjusting textgrids from MFA with the standard Praat interface (Boersma, 2024).

Phonetic annotations include only target words, not complete carrier sentence. The validation sample has up to 40 L1 + 40 L2 tokens per word, but in the extension sample this is reduced to 20 each, as pilot evaluation established this to be sufficient. An error tier was added to L2 speakers' textgrids, using a simple coding scheme to mark when any of consonant, vowel quality, quantity, and/or stress placement errors were present in the target word. Most prominent among errors in vowel quality was a blending of the distinct vowel pairs i (L1 [ɪ]) and í (L1 [i]), as well as o (L1 [ɔ]) and ó (L1 [ou]), possibly explained by their orthographic similarity.

Segment	N	Same	10%	25ms	Error
L1-Ons	1617	64%	70%	87%	27ms
L2-Ons	931	69%	77%	91%	29ms
L1-V	1727	48%	62%	79%	31ms
L2-V	980	64%	76%	87%	29ms
L1-C	1727	57%	70%	83%	29ms
L2-C	980	66%	75%	85%	37ms
L1-Ratio	1727	42%	67%	–	19%
L2-Ratio	980	56%	75%	–	20%

Table 1: MFA accuracy for Onset, stressed Vowel, and post-vowel Consonant segment durations, and resulting V/(V+C) Ratio. Columns are: Number of tokens; percent of tokens where MFA's duration/ratio is the Same, within 10%, or within 25ms of gold; and average magnitude of MFA Errors.

4 Evaluations

4.1 MFA Alignment Accuracy

First, automatic (MFA) phone alignments are compared to manual (gold) annotations (Table 1). MFA output has accurate durations for half to 2/3 of relevant segments, and of the rest, annotators' adjustments are on average around 30ms. MFA inaccuracies affect the V/(V+C) ratio for roughly half of tokens, on average by 19-20% of the actual ratio.

4.2 Quantity classification

For a first look at acoustic correlates of the quantity contrast, K-nearest-neighbour (K=1,3,5,10,20) and linear regression classifiers were trained to predict vowels' phonological length, using the following features extracted from gold (manual) annotations and MFA (automated) forced alignments: V/(V+C) **Ratio**, and segment durations **OnsDur**, **VDur**, **CDur**, and **WordDur** of respectively the target syllable Onset, Vowel, following Consonant(s), and whole Word. Classifiers use 5-fold cross validation, or leave-one-out cross validation for samples under 100 tokens. In §4 only the most informative feature sets are reported, using 5-nearest-neighbour classifiers which were typical of overall results.

In Table 2, a classifier for **All** tokens in the dataset has mediocre accuracy (L1 gold: 75%) using the V/(V+C) Ratio, with limited improvement from other available features. §4.3-4.6 therefore use linguistically restricted subsets of the data, aiming to isolate factors that moderate quantity cues.

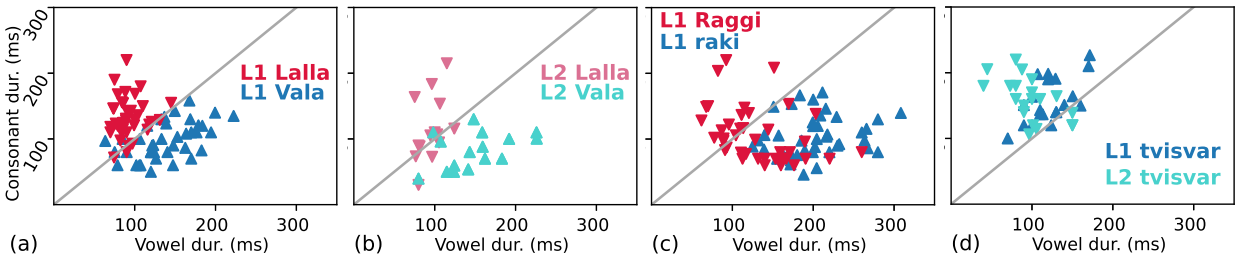

Figure 1: Stressed vowel and following consonant durations in *Lalla, Vala, Raggi, raki,* and *tvisvar*.

Sample	Features	L1-Gold	L1-MFA	L2-Gold	L2-MFA
All	Ratio	75%	74%	69%	69%
All	VDur	68%	71%	61%	60%
All	OnsDur, VDur, CDur	79%	79%	70%	70%
[C]ALa	Ratio	98%	93%	91%	84%
[C]ALa	VDur	84%	80%	66%	62%
[C]ALa	VDur, Cdur	99%	95%	91%	88%
*ALa	Ratio	94%	95%	81%	89%
*ALa	OnsDur, VDur, CDur	96%	96%	87%	90%
*AKi	Ratio	66%	68%	71%	68%
*AKi	VDur	67%	71%	78%	69%
*AKi	VDur, Cdur, WordDur	74%	76%	67%	72%
haus-	Ratio	100%	98%	76%	74%
Diphthong	Ratio	98%	97%	74%	76%

Table 2: Vowel length KNN classifier accuracy for L1 and L2 speech, with features computed from gold (manual) annotations and MFA alignments. Samples consist of: **All** 72 words of the dataset; **[C]ALa**: *dala, gala, tala, balla, galla, kalla, palla*; ***ALa**: the previous class plus *ala, fala, vala, dvala, svala, lalla, malla*; ***AKi**: *aki, aggi, baki, baggi, taki, kaggi, raki, raggi, þaki, blaki, maki, maggi*; **haus-**: *hausinn, hausnum*; **Diphthong**: *ása, ásta, hausinn, hausnum, jónas, jónsson*.

4.3 *-ala, -alla*

Results for **[C]ALa** in Table 2 examine two-syllable words of a plosive followed by [aːla] or [alːa], parallel to Pind (1995). Ratio is almost completely sufficient to distinguish L1 quantity (gold: 98% accuracy), while as expected, vowel duration (VDur) alone is not. However, VDur and CDur jointly may be slightly more useful than Ratio, especially with MFA features. ***ALa**, with more syllable onset types, is harder to classify by Ratio, but providing onset duration as a moderating factor may make up some of the difference, especially for L2 speakers. In all cases L2 speech was not classified as accurately as L1; examples of short (*Lalla*, personal name) and long (*Vala*, personal name) vowels in Figures 1a-b illustrate how short and long vowel cues overlap less for L1.

4.4 *-aki, -aggi*

***AKi** in Table 2 finds far worse ability to discriminate vowel quantity than either Pind (1999)'s 94% (L1) classification accuracy for similar words with only single plosive onsets, or to our ***ALa** sample with varied onsets. Figure 1c gives an example of L1 speech for minimal pair *raki* [raːcɪ] 'humidity', *Raggi* [racːɪ] (personal name), clearly not separable by the features that were sufficient for ***ALa**. For L1 but not L2, whole token duration is somewhat useful; this feature can reflect local speech rate and aspects of onset consonants.

4.5 Consonant cluster exceptions

In *tvisvar* 'twice' (Figure 1d), long [ɪː] precedes an 'exceptional' cluster [sv]. L1 and L2 consonant cluster durations are all around 100-225ms, but L1 vowel durations (most 100-200ms) are notably longer than L2 (many under 100ms, few over 125ms). Reading *i* in *tvisvar* as a short vowel may show partially successful L2 acquisition of a vowel quantity system, but failure to incorporate nuances.

4.6 Diphthongs

[œiː] and [œi] in the words *hausinn, hausnum* ('the head', nominative and dative respectively) are distinct for L1 speakers, but not L2, who tend to insufficiently reduce diphthong duration in *hausnum*. This is unsurprising, as contrastive 'short' diphthongs are typologically rare. The observation generalises to **Diphthongs** (Table 2) with more vowel qualities and contexts, indicating promise for an area where CAPT may provide valuable feedback.

5 Discussion

At a high level, RQ1 is answered positively, as the conventional ratio proves to be an informative and interpretable feature, and more useful than absolute vowel duration alone. More specifically, for *-ala, -alla* words, expectations from a controlled study were strongly upheld in our crowdsourced data. For *-aki, -aggi*, aggregated data also would seem to match expectations, but a substantial proportion of individual tokens occupy an ambiguous region, at least in all currently examined feature spaces.

Regarding RQ2, quantity can be classified from the Ratio feature, but long and short vowels are not always well separable, and absolute durations of vowel and consonant carry some useful information beyond the ratio. Location of a best threshold for any features also varies based on several other factors. In some cases, factors are identified and controlled for, with good to excellent classifier performance. In other cases this work is ongoing, and a general-purpose solution remains to be developed; it could require phoneme identity labels, representations of syllable and word structure, prosodic environment, spectral features, etc. Qualitatively, during annotation we had observed noticeable length errors in some of the same L2 samples (e.g. *tvisvar, hausnum*) where the measured features indicated loss of contrast for L2 speakers as compared to L1, which is an encouraging sign that the features can capture perceptually important dimensions of contrast.

Addressing RQ3, MFA frequently mismeasures segments in this corpus by around one-third of the true duration, although the particular values for all MFA measures arise from a specific acoustic model and do not generalise to others. The relevant interpretation is that typical applications of MFA, like ours obtaining decent word alignments from 20 hours of in-domain training speech, cannot be relied on for the accuracy desired by primary descriptive research in phonetic segments. MFA preprocessing may also introduce bias in the gold annotations, which would not critically affect CAPT development and is well worth the saved time over full manual transcription, but true inaccuracy of MFA may be underestimated.

Despite considerable room for improvement, alignment errors had small impacts (RQ4) on classifiers' ability to distinguish short and long vowels. The relative utility of various features also appears similar whether using manual or automated data.

5.1 Contributions

We freely release our annotations, whose audio and metadata is already public. The data is available at *https://github.com/catiR/length-contrast-data-isl* accompanied by an online platform for visualisations/analyses as in §4. This is the most accessible data on L1 Icelandic vowel length, and the first L2 data. Preliminary analysis reveals L2 acquisition of a quantity contrast to an extent, but also some systematic challenges.

Towards CAPT software development, MFA and temporal features derived from it are identified as an adequate starting point to (i) characterise distinctiveness or ambiguity of typical pronunciations across phonological quantity contrasts; and (ii) classify apparent quantity of L2 pronunciations, and alert learners to errors if their pronunciation is unambiguously not what it should be.

Acknowledgments

We would like to thank Branislav Bédi, Gunnar Thor Örnólfsson, Luke O'Brien, Marc Daníel Skipstað Volhardt, Staffan Hedström, Þorsteinn Daði Gunnarsson, and all of our colleagues involved in pedagogic development of Icelandic computer assisted language learning and collecting L1 and L2 speech.

This work was supported by The Icelandic Centre for Research (RANNÍS), under the Icelandic Student Innovation Fund project *Speech corpus of L1 and L2 Icelandic vowel length*, Grant 2412155-1101.

References

Kristján Árnason. 1998. Vowel shortness in Icelandic. *Phonology and morphology of the Germanic languages*, pages 3–25.

Kristján Árnason. 2011. *The phonology of Icelandic and Faroese*. Oxford University Press.

Birna Arnbjörnsdóttir, Kolbrún Friðriksdóttir, and Branislav Bédi. 2020. Icelandic online: twenty years of development, evaluation, and expansion of an LMOOC. *CALL for widening participation: short papers from EUROCALL*, pages 13–19.

Branislav Bedi, Jane O'Toole, and Monica Ward. 2024. Resourceful approaches in call for less-commonly taught languages (LCTLs): Case studies on Icelandic, Irish, and Nawat. *EuroCALL 2023: CALL for all Languages*.

Paul Boersma. 2024. Praat: doing phonetics by computer [Computer program]. *http://www.praat.org/*.

Branislav Bédi. 2022. Development of online tools supporting the learning of Icelandic as a foreign and second language. In Branislav Bédi, Halldóra J. Þorláksdóttir, and Kolbrún Friðriksdóttir, editors, *Tungumál í samhengi / Perspectives on Language and Context*. Reykjavík: Stofnun Vigdísar Finnbogadóttur í erlendum tungumálum.

Radu Craioveanu. 2023. *Weighing Preaspiration: Phonetics, Phonology, & Typology of a Laryngeal Phenomenon*. Ph.D. thesis, University of Toronto (Canada).

Stefán Einarsson. 1927. *Beiträge zur Phonetik der isländischen Sprache*. AW Brøggers.

Stefán Einarsson. 1945. *Icelandic: Grammar, texts, glossary*. The Johns Hopkins Press.

Marcin Fortuna. 2016. Icelandic post-lexical syllabification and vowel length in CVCV phonology. *The Linguistic Review*, 33(2):239–275.

Edmund Gussmann. 2011. Getting your head around: the vowel system of Modern Icelandic. *Folia Scandinavica Posnaniensia*, 12:71–91.

Staffan Hedström, Judy Y Fong, Ragnheiður Þórhallsdóttir, David Erik Mollberg, Smári Freyr Guðmundsson, Ólafur Helgi Jónsson, Sunneva Þorsteinsdóttir, Eydís Huld Magnúsdóttir, and Jon Gudnason. 2021. Samromur queries 21.12. CLARIN-IS.

Staffan Hedström, Judy Y. Fong, Ragnheiður Þórhallsdóttir, David Erik Mollberg, Smári Freyr Guðmundsson, Ólafur Helgi Jónsson, Sunneva Þorsteinsdóttir, Eydís Huld Magnúsdóttir, and Jon Gudnason. 2022a. Samromur unverified 22.07. CLARIN-IS.

Staffan Hedström, Judy Y. Fong, Ragnheiður Þórhallsdóttir, David Erik Mollberg, Thomas Mestrou, Smári Freyr Guðmundsson, Ólafur Helgi Jónsson, Sunneva Þorsteinsdóttir, Eydís Huld Magnúsdóttir, Caitlin Laura Richter, Ragnar Pálsson, and Jon Gudnason. 2022b. Samromur L2 22.09. CLARIN-IS.

Ari Páll Kristinsson. 1988. *The pronunciation of modern Icelandic: a brief course for foreign students*. Málvísindastofnun Háskóla Íslands.

Ari Páll Kristinsson, Friðrik Magnússon, Margrét Pálsdóttir, and Sigrún Þorgeirsdótir. 1985. Um andstæðuáherslu í íslensku. *Íslenskt mál og almenn maálfræði*, 7:7–47.

Robert McAllister, James E Flege, and Thorsten Piske. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of phonetics*, 30(2):229–258.

Michael McAuliffe, Michaela Socolof, Sarah Mihuc, Michael Wagner, and Morgan Sonderegger. 2017. Montreal forced aligner: Trainable text-speech alignment using kaldi. In *Interspeech*, volume 2017, pages 498–502.

David Erik Mollberg, Ólafur Helgi Jónsson, Sunneva Þorsteinsdóttir, Jóhanna Vigdís Guðmundsdóttir, Steinþór Steingrímsson, Eydís Huld Magnúsdóttir, Judy Y Fong, Michal Borsky, and Jon Gudnason. 2021. Samromur 21.05. CLARIN-IS.

Anna Björk Nikulásdóttir and Jón Guðnason. 2017. General pronunciation dictionary for ASR. CLARIN-IS.

Jörgen Pind. 1995. Speaking rate, voice-onset time, and quantity: The search for higher-order invariants for two Icelandic speech cues. *Perception & Psychophysics*, 57(3):291–304.

Jörgen Pind. 1999. Speech segment durations and quantity in Icelandic. *The Journal of the Acoustical Society of America*, 106(2):1045–1053.

Höskuldur Þráinsson. 1994. Icelandic. In *The Germanic Languages*, pages 142–189. Routledge.

Appendix A. Detailed annotation contents

Word	L1 tokens	L2 tokens	Carriers
ala	2	3	4
dala	40		14
dvala	40		22
fala	1	1	2
gala	5		2
svala	40	4	17
tala	40	40	47
vala	40	16	41
aki	9		3
baki	40	40	65
blaki	16	2	4
maki	24	4	10
raki	42	1	13
taki		19	15
þaki		9	6
breki		40	30
dreki	40	5	18
leki	66	1	17
speki	40	2	17
veki	25		7
balla	1		1
galla	10	5	4
kalla	33	2	15
lalla	40	14	36
malla	37	4	14
palla	40	9	26
aggi	19		6
baggi	26		8
kaggi	11		1
maggi	40	25	34
raggi	41	11	25
beggi	29	7	11
eggi	40	7	23
leggi	41	3	20
skeggi	40	2	16
veggi	40	6	28

Table 3: Counts of L1 and L2 tokens, and unique carrier sentences, for words in the validation sample. While even distribution was a guiding principle, the contents are necessarily a compromise between balance and availability of data.

Word	L1 tokens	L2 tokens	Carriers
ása	20	20	16
bera	21	20	28
betri	20	23	15
brosir	20	20	25
fara	21	20	15
færa	20	18	29
færi	20	20	38
hausinn	20	20	14
jónas	20	20	28
katrín	20	20	31
kisa	20	12	15
koma	20	20	33
leyfa	20	20	21
muna	20	21	22
nema	20	20	15
sama	20	20	20
sækja	20	20	39
sömu	20	20	16
tvisvar	20	20	28
vinur	20	20	35
ásta	20	19	18
farðu	21	18	23
fossinn	20	15	13
færði	20	20	34
hausnum	20	22	18
herra	20	20	21
jónsson	20	20	29
leyfðu	26	20	23
mamma	20	20	15
missa	20	20	37
mömmu	20	20	15
nærri	20	20	34
snemma	20	20	17
sunna	20	20	21
tommi	20	20	32
vinnur	20	10	27

Table 4: Counts of L1 and L2 tokens, and unique carrier sentences, for words in the extension sample. Long vowels are in the upper section of the table and short vowels in the lower.