Incorporating tone in the calculation of phonotactic probability

James P. Kirby
University of Edinburgh
School of Philosophy, Psychology, and Language Sciences
j.kirby@ed.ac.uk

Abstract

This paper investigates how the ordering of tone relative to the segmental string influences the calculation of phonotactic probability. Trigram and recurrent neural network models were trained on syllable lexicons of four Asian syllable-tone languages (Mandarin, Thai, Vietnamese, and Cantonese) in which tone was treated as a segment occurring in different positions in the string. For trigram models, the optimal permutation interacted with language, while neural network models were relatively unaffected by tone position in all languages. In addition to providing a baseline for future evaluation, these results suggest that phonotactic probability is robust to choices of how tone is ordered with respect to other elements in the syllable.

1 Introduction

The phonotactic probability of a string is an important quantity in several areas of linguistic research, including language acquisition, wordlikeness, word segmentation, and speech production and perception (Bailey and Hahn, 2001; Daland and Pierrehumbert, 2011; Storkel and Lee, 2011; Vitevitch and Luce, 1999). When the language of interest is a tone language, the question arises of how tone should be incorporated into the probability calculation. As phonotactic probability is frequently computed based on some type of n-gram model, this means deciding on which segment(s) the probability of a tone should be conditioned. For instance, using a bigram model, one might compute the probability of the Mandarin syllable *fāng* as $P(\text{a}|\text{f}) \times P(\text{ŋ}|\text{a}) \times P(\text{tone 1}|\text{ŋ})$, but could just as well consider $P(\text{tone 1}|\text{f}) \times P(\text{a}|1) \times P(\text{ŋ}|\text{a})$, or any other conceivable permutation of tone and segments.

While this issue is occasionally remarked on (e.g. Newman et al., 2011: 246), there remains no widespread consensus in practice. Choice of ordering is sometimes justified based on segment-tone co-occurrence restrictions in the language under study (Myers and Tsay, 2005), but is often presented without justification (Kirby and Yu, 2007; Yang et al., 2018), and in some cases tone is simply ignored (Gong, 2017). When the space of possibilities is considered, researchers generally select the permutation which maximizes model fit to some external data, such as participant judgments of phonological distance (Do and Lai, 2021a) or wordlikeness (Do and Lai, 2021b).

Although extrinsic evaluation is in some sense a gold standard, intrinsic metrics of model fit can also be informative, in part because extrinsic metrics are not always robust across data sets. For instance, participant wordlikeness judgments can vary considerably based on the particulars of the experimental design (Myers and Tsay, 2005; Shademan, 2006; Vitevitch and Luce, 1999), so the treatment of tone that produces a best-fit model for one dataset may not do so for another. The lexicon of a given language is much more internally stable in terms of how segments and tones are distributed, so intrinsic evaluation may provide a useful baseline for reasoning about the treatment of tone relative to segments both within and across languages.

This short paper considers a simple information-theoretic motivation for selecting a permutation: all else being equal, we should prefer a model that maximizes the probability of the lexicon (i.e., minimizes the cross-entropy loss), because this will be the model that by definition does the best job of capturing the phonotactic regularities of the lexicon (Cherry

Proceedings of the Seventeenth SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology,pages 32–38
August 5, 2021. ©

et al., 1953; Goldsmith, 2002; Pimentel et al., 2020). By treating tone as another phone in the segmental string, we can see whether and to what degree this choice has an effect on the overall entropy of the lexicon.

Intuitively, *any* model that can take into account phonotactic constraints will result in a reduction in entropy. Thus, even an n-gram model with a sufficiently large context window should in principle be able model segment-tone co-occurrences at the syllable level. However, tone languages differ with respect to tone-segment co-occurrence restrictions (see Sec. 2). If a relevant constraint primarily targets syllable onsets, for instance, placing the tonal "segment" in immediate proximity to the onset will increase the probability of the string, even relative to a model capable of capturing the dependency at a longer distance.

2 Languages

Four syllable-tone languages were selected for this study: Mandarin Chinese, Cantonese, Vietnamese and Thai. They are partially a convenience sample in that the necessary lexical resources were readily available, but also have some useful similarities: all share a similar syllable structure template and have five or six tones. However, the four languages vary in terms of their segment-tone co-occurrence restrictions, as detailed below.

In all cases, the lexicon was defined as the set of unique syllable shapes in each language. For consistency, the syllable template in all four languages is considered to be $(C_1)(C_2)V(C)T$, with variable positioning of T. Offglides were treated as codas in all languages. The syllable lexicons for all four languages are provided in the supplementary materials (`http://doi.org/10.17605/OSF.IO/NA5FB`).

Mandarin (`cmn`) The Mandarin syllabary consists of 1,226 syllables based on list of attested readings of the 13,060 BIG5 characters from Tsai (2000), phonetized using the phonological system of Duanmu (2007). This representation encodes 22 onsets, 3 medials (/j ɥ w/), 6 nuclei, 4 codas and 5 tones (including the neutral tone). In Mandarin, unaspirated obstruent onsets rarely appear with mid-rising tone (MC *yang ping*), and sonorant onsets rarely occur with the high-level tone (MC *yin ping*). Obstruents never occur as codas.

Thai (`tha`) A Thai lexicon of 4,133 unique syllables was created based on the dictionary of Haas (1964) which contains around 19,000 entries and 47,000 syllables. The phonemic representation encodes 20 onsets, 3 medials /w l r/, 21 nuclei (vowel length being contrastive in Thai), 8 codas and 5 tones. In Thai, high tone is rare/unattested following unaspirated and voiced onsets, but there is also statistical evidence for a restriction on rising tones with these onsets (Perkins, 2013). In syllables with an obstruent coda (/p t k/), only high, low, or falling tones occur, depending on length of the nuclear vowel (Morén and Zsiga, 2006).

Vietnamese (`vie`) The Vietnamese lexicon of 8,128 syllables was derived from a freely available dictionary of around 74,000 words (Đức, 2004), phonetized using a spelling pronunciation (Kirby, 2008). The resulting representation encodes 24 onsets, 1 medial (/w/), 14 nuclei, 8 codas and 6 tones. Vietnamese syllables ending in obstruents /p t k/ are restricted to just one of two tones.

Cantonese (`yue`) The Cantonese syllabary consists of the 1,884 unique syllables in the Chinese Character Database (Kwan et al., 2003), encoded using the *jyutping* system. This representation distinguishes 22 onsets, 1 medial (/w/), 11 nuclei, 5 codas and 6 tones. In Cantonese, unaspirated initials do not occur in syllables with low-falling tones, and aspirated initials do not occur with the low tone. Syllables ending with /p t k/ are restricted to one of the three "entering" tones (Yue-Hashimoto, 1972).

3 Methods

Two classes of character-level language models (LMs) were considered: simple n-gram models and recurrent neural networks (Mikolov et al., 2010). In an n-gram model, the probability of a string is proportional to the conditional probabilities of the component n-grams:

$$P(x_i|x_1^{i-1}) \approx P(x_i|x_{i-n+1}^{i-1}) \qquad (1)$$

The degree of context taken into account is thus determined by the value chosen for n.

In a recurrent neural network (RNN), the next character in a sequence is predicting using the current character and the previous hidden state. At each step t, the network retrieves an embedding for the current input x_t and combines it with the hidden layer from the previous step to compute a new hidden layer h_t:

$$h_t = g(Uh_{t-1} + Wx_t) \tag{2}$$

where W is the weight matrix for the current time step, U the weight matrix for the previous time step, and g is an appropriate nonlinear activation function. This hidden layer h_t is then used to generate an output layer y_t, which is passed through a softmax layer to generate a probability distribution over the entire vocabulary. The probability of a sequence $x_1, x_2 \ldots x_z$ is then just the product of the probabilities of each character in the sequence:

$$P(x_1, x_2 \ldots x_z) = \prod_{i=1}^{z} y_i \tag{3}$$

The incorporation of the recurrent connection as part of the hidden layer allows RNNs to avoid the problem of limited context inherent in n-gram models, because the hidden state embodies (some type of) information about *all* of the preceding characters in the string. Although RNNs cannot capture arbitrarily long-distance dependencies, this is unlikely to make a difference for the relatively short distances involved in phonotactic modeling.

Trigram models were built using the SRILM toolkit (Stolcke, 2002), with maximum likelihood estimates smoothed using interpolated Witten-Bell discounting (Witten and Bell, 1991). RNN LMs were built using PyTorch (Paszke et al., 2019), based on an implementation by Mayer and Nelson (2020). The results reported here make use of simple recurrent networks (Elman, 1990), but similar results were obtained using an LSTM layer (Hochreiter and Schmidhuber, 1997).

3.1 Procedure

The syllables in each lexicon were arranged in 5 distinct permutations: tone following the coda (T|C), nucleus (T|N), medial (T|M), onset (T|O) and with tone as the initial segment in the syllable (T|#). As many syllables in these languages lack onsets, medials, and/or codas, a sizable number of the resulting strings were identical across permutations. Both smoothed trigram and simple RNN LMs were then fit to each permuted lexicon 10 times, with random 80/20 train/dev splits (other splits produced similar results). For each run, the perplexity of the language model on the dev set $D = x_1x_2 \ldots x_N$ (i.e., the exponentiated cross-entropy[1]) was recorded:

$$\begin{aligned} PPL(D) &= b^{H(D)} && (4) \\ &= b^{-\frac{1}{N} log_b P(x_1 x_2 \ldots x_N)} && (5) \end{aligned}$$

4 Results

For brevity, only the main findings are summarized here; the full results are available as part of the online supplementary materials (`http://doi.org/10.17605/OSF.IO/NA5FB`).

Table 1 show the orderings which minimized perplexity for each method and language, averaged over 10 runs. Table 2 shows the average perplexity over all permutations for a given language and method.

method	lexicon	order	PPL
3-gram	`cmn`	T|C	4.91 (0.06)
	`tha`	T|M	7.34 (0.12)
	`vie`	T|C	7.35 (0.03)
	`yue`	T|M	5.84 (0.09)
RNN	`cmn`	T|M	4.01 (0.08)
	`tha`	T|M	5.20 (0.04)
	`vie`	T|M	5.16 (0.02)
	`yue`	T|#	4.37 (0.05)

Table 1: Orders which produced the lowest perplexities averaged over 10 runs (means and standard deviations).

Differences between orderings were then assessed visually, aided by simple analyses of variance. For the trigram LMs, perplexity was lowest in Mandarin when tones followed codas, while differences in perplexity between other orderings were negligible. For Thai, Vietnamese, and Cantonese, all orderings were roughly comparable except for when tone was ordered as the first segment in the syllable (T|#), which increased perplexity by up to 1 over the mean of the other orderings. For Thai, the ordering T|M resulted in significantly lower perplexities compared to all other

[1] Equivalently, we may think of $PPL(D)$ as the inverse probability of the set of syllables D, normalized for the number of phonemes.

	cmn	tha	vie	yue
3-gram	5.15 (0.17)	7.76 (0.4)	7.49 (0.27)	5.98 (0.18)
RNN	4.01 (0.07)	5.28 (0.05)	5.18 (0.03)	4.42 (0.07)

Table 2: Mean and standard deviation of perplexity across all permutations by lexicon and language model.

permutations. For the RNN LMs, although T|M was the numerically optimal ordering for three out of the four languages, in practical terms permutation had no effect on perplexity, with numerical differences of no greater than 0.1 (see Table 2).

5 Discussion

Consistent with other recent work in computational phonotactics (e.g. Mayer and Nelson, 2020; Mirea and Bicknell, 2019; Pimentel et al., 2020), the neural network models outperformed the trigram baselines by a considerable margin (a reduction in average perplexity of up to 2.5, depending on language). Neural network models were also much less sensitive to the linear position of tone relative to other elements in the segmental string (cf. Do and Lai, 2021b), no doubt due to the fact that the ability of the RNNs to model co-occurrence tendencies within the syllable is not constrained by context in the way that *n*-gram models are.

Perhaps as a result, however, the RNN models reveal little about the nature of segment-tone co-occurrence restrictions in any of the languages investigated. In this regard, the trigram models, while clearly less optimal in a global sense, are still informative. The fact that the ordering T|# was significantly worse under the trigram model for Cantonese, Vietnamese and Thai but not Mandarin can be explained (or predicted) by the fact that of the four languages, only Mandarin does not permit obstruent codas, and consequently has no coda-tone co-occurrence restrictions (indeed, the four primary tones of Mandarin occur with more or less equal type frequency). In the other three languages, syllables with obstruent codas can only bear a restricted set of tones, and in a trigram model, this dependency is not modeled when tone is prepended to the syllable, since this means it will frequently, though not always, fall outside the window visible to the language model. Even a model with a large enough context window to capture such dependencies will assign the lexicon a higher perplexity when structured in this way.

The finding that the T|M ordering is always optimal in Thai (and by a larger margin than in the other languages) is presumably due to the fact that the distribution of the medials /w l r/ is severely restricted in this language, occurring only after /p $\mathrm{p^h}$ t $\mathrm{t^h}$ k $\mathrm{k^h}$ f/. The distribution of tones after onset-medial clusters is inherently more constrained and therefore more predictable. A similar restriction holds in Cantonese, albeit to a lesser degree (the medial /w/ only occurs with onsets /k/ and /$\mathrm{k^h}$/).

5.1 Shortcomings and extensions

This work did not explore representations based on phonological features, given that their incorporation has failed to provide evaluative improvements in other studies of computational phonotactics (Mayer and Nelson, 2020; Mirea and Bicknell, 2019; Pimentel et al., 2020). However, feature-based approaches can be both theoretically insightful and may even prove necessary for other quantifications, such as the measure of phonological distance where tone is involved (Do and Lai, 2021a).

The present study has focused on a small sample of structurally and typologically similar languages. All have relatively simple syllable structures in which one and only one tone is associated with each syllable. Not all tone languages share these properties, however. In so-called "word-tone" languages, such as Japanese or Shanghainese, the surface tone with which a given syllable is realized is frequently not lexically specified. In other languages, such as Yoloxóchitl Mixtec (DiCanio et al., 2014), tonal specification may be tied to sub-syllabic units, such as the mora. Finally, data from many other languages, such as Kukuya (Hyman, 1987), make it clear that in at least in some cases tones can only be treated in terms of abstract melodies, which do not have a consistent association to syllables, moras, or vowels (Goldsmith, 1976). In these and many other cases, careful consideration of the theoretical motivations justifying a particular representation are required before it makes sense to consider ordering effects.

However, to the extent that it is possible to generate a segmental representation of a tone language in which surface tones are indicated, what the present work suggests is that the precise ordering of the tonal symbols with respect to other symbols in the string is unlikely to have a significant impact on phonotactic probability. This follows from two assumption (or constraints): first, that the set of symbols used to indicate tones is distinct from those used to indicate the vowels and consonants; and second, that one and only one such tone symbol appears per string domain (here, the syllable). If these two constraints hold, the complexity of the syllable template should in general have a greater impact on the entropy of the string set than the position of the tone symbol, although the number of unique tone symbols relative to the number of segmental symbols may also have an effect. According to Maddieson (2013) and Easterday (2019), languages with complex syllable structures (defined as those permitting fairly free combinations of two or more consonants in the position before a vowel, and/or two or more consonants in the position after the vowel) rarely have complex tone systems, or indeed tone systems at all, so this is unlikely to be an issue for most tone languages.

One possibility the present work did not address is whether it is even necessary, or desirable, to include tone in phonotactic probability calculations in the first place. The probability of the lexicon of a tonal language would surely change if tone is ignored, but whether listeners' judgments of a sequence as well- or ill-formed is better predicted by a model that takes tone into account vs. one that does not is an empirical question (but see Kirby and Yu, 2007; Do and Lai, 2021b for some evidence that it may not). Similarly, for research questions focused on tone sandhis, or on the distributions of the tonal sequences themselves (tonotactics), the relevant computations will be restricted to the tonal tier in the first instance, and ordering with respect to segments may simply not be relevant (but see Goldsmith and Riggle, 2012).

Finally, the present study has focused on the lexical representation of tone, but in many languages tone primarily serves a morphological function. The SIGMORPHON 2020 Task 0 shared challenge (Vylomova et al., 2020) included inflection data from several tonal Oto-Manguean languages in which tone was orthographically encoded in different ways via string diacritics. While the authors noted the existence these differences, it is unclear whether and to what extent the different representations of tones affected system performance. Similarly, the potential impact of tone ordering relative to other elements in the string has yet to be systematically investigated in this setting.

6 Conclusion

This paper has assessed how different permutations of tone and segments affects the perplexity of the lexicon in four syllable-tone languages using two types of phonotactic language models, an interpolated trigram model and a simple recurrent neural network. The perplexities assigned by the neural network models were essentially unaffected by different choices of ordering; while the trigram model was more sensitive to permutations of tone and segments, the effects on perplexity remained minimal. In addition to providing a baseline for future evaluation, these results suggest that the phonotactic probability of a syllable is relatively robust to choice of how tone is ordered with respect to other elements in the string, especially when using a model capable of encoding dependencies across the entire syllable.

Acknowledgments

This work was supported in part by the ERC EVOTONE project (grant no. 758605).

References

Todd Bailey and Ulrike Hahn. 2001. Determinants of wordlikeness: phonotactics or lexical neighborhoods? *Journal of Memory and Language*, 44:568–591.

E. Colin Cherry, Morris Halle, and Roman Jakobson. 1953. Toward the logical description of languages in their phonemic aspect. *Language*, 29(1):34–46.

Robert Daland and Janet B. Pierrehumbert. 2011. Learning diphone-based segmentation. *Cognitive Science*, 35(1):119–155.

Christian DiCanio, Jonathan D Amith, and Rey Castillo García. 2014. The phonetics of moraic alignment in yoloxóchitl mixtec. In *Proceedings of the 4th International Symposium on Tonal Aspects of Languages (TAL-2014)*, pages 203–210.

Youngah Do and Ryan Ka Yau Lai. 2021a. Accounting for lexical tones when modeling phonological distance. *Language*, 97(1):e39–e67.

Youngah Do and Ryan Ka Yau Lai. 2021b. Incorporating tone in the modelling of wordlikeness judgements. *Phonology*, 37:577–615.

San Duanmu. 2007. *The phonology of standard Chinese*, 2nd edition. Oxford University Press, Oxford.

Shelece Easterday. 2019. *Highly complex syllable structure: A typological and diachronic study*. Studies in Laboratory Phonology. Language Science Press.

Jeffrey L. Elman. 1990. Finding structure in time. *Cognitive Science*, 14(2):179–211.

John Goldsmith. 1976. *Autosegmental Phonology*. Ph.D. thesis, MIT. [Published by Garland Press, New York, 1979.].

John Goldsmith. 2002. Phonology as information minimization. *Phonological Studies*, 5:21–46.

John Goldsmith and Jason Riggle. 2012. Information theoretic approaches to phonological structure: the case of Finnish vowel harmony. *Natural Language and Linguistic Theory*, 30:859–896.

Donald Shuxiao Gong. 2017. Grammaticality and lexical statistics in Chinese unnatural phonotactics. *UCL Working Papers in Linguistics*, 17:1–23.

Mary R. Haas. 1964. *Thai-English student's dictionary*. Stanford University Press, Stanford.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Larry M. Hyman. 1987. Prosodic domains in Kukuya. *Natural Language & Linguistic Theory*, 5(3):311–333.

James P. Kirby. 2008. vPhon: a Vietnamese phonetizer (version 2.1.1) [computer program]. https://github.com/kirbyj/vPhon.

James P. Kirby and Alan C. L. Yu. 2007. Lexical and phonotactic effects on wordlikeness judgments in Cantonese. In *Proceedings of the 16th International Conference of the Phonetic Sciences*, pages 1389–1392, Saarbrücken.

Tze-Wan Kwan, Wai-Sang Tang, Tze-Ming Chiu, Lei-Yin Wong, Denise Wong, and Li Zhong. 2003. Chinese character database with word-formations phonologically disambiguated according to the Cantonese dialect. http://humanum.arts.cuhk.edu.hk/Lexis/lexican/. Accessed 9 February 2007.

Ian Maddieson. 2013. Tone. In Matthew S. Dryer and Martin Haspelmath, editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology.

Connor Mayer and Max Nelson. 2020. Phonotactic learning with neural language models. *Proceedings of the Society for Computation in Linguistics*, 3:16.

Tomáš Mikolov, Martin Karafiát, Lukáš Burget, Jan Černocký, and Sanjeev Khudanpur. 2010. Recurrent neural network based language model. In *Proc. INTERSPEECH 2010*, page 1045–1048.

Nicole Mirea and Klinton Bicknell. 2019. Using LSTMs to assess the obligatoriness of phonological distinctive features for phonotactic learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1595–1605. Association for Computational Linguistics.

Bruce Morén and Elizabeth Zsiga. 2006. The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory*, 24(1):113–178.

James Myers and Jane Tsay. 2005. The processing of phonological acceptability judgements. In *Proc. Symposium on 90-92 NSC Projects*, Taipei.

Ellen Hamilton Newman, Twila Tardif, Jingyuan Huang, and Hua Shu. 2011. Phonemes matter: The role of phoneme-level awareness in emergent Chinese readers. *Journal of Experimental Child Psychology*, 108(2):242–259.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc.

Jeremy Perkins. 2013. *Consonant-tone interaction in Thai*. Ph.D. thesis, Rutgers, The State University of New Jersey.

Tiago Pimentel, Brian Roark, and Ryan Cotterell. 2020. Phonotactic complexity and its trade-offs. *Transactions of the Association for Computational Linguistics*, 8:1–18.

Shabnam Shademan. 2006. Is phonotactic knowledge grammatical knowledge? In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, pages 371–379. Cascadilla Proceedings Project.

Andreas Stolcke. 2002. SRILM – an extensible language modeling toolkit. In *Proc. Intl. Conf. on Spoken Language Processing Vol. 2*, pages 901–904, Denver.

Holly L. Storkel and Su-Yeon Lee. 2011. The independent effects of phonotactic probability and neighbourhood density on lexical acquisition by preschool children. *Language and Cognitive Processes*, 26(2):191–211.

Chih-Hao Tsai. 2000. Mandarin syllable frequency counts for Chinese characters. http://technology.chtsai.org/syllable/. Accessed 10 March 2021.

Michael S. Vitevitch and Paul A. Luce. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40:374–408.

Ekaterina Vylomova, Jennifer White, Elizabeth Salesky, Sabrina J. Mielke, Shijie Wu, Edoardo Maria Ponti, Rowan Hall Maudslay, Ran Zmigrod, Josef Valvoda, Svetlana Toldova, and et al. 2020. Sigmorphon 2020 shared task 0: Typologically diverse morphological inflection. In *Proceedings of the 17th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, page 1–39. Association for Computational Linguistics.

Ian H. Witten and Timothy C. Bell. 1991. The zero-frequency problem: estimating the probabilities of novel events in adaptive text compression. *IEEE Transactions on Information Theory*, 37(4):1085–1094.

Shiying Yang, Chelsea Sanker, and Uriel Cohen Priva. 2018. The organization of lexicons: a cross-linguistic analysis of monosyllabic words. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2018*, page 164–173.

Anne O. Yue-Hashimoto. 1972. *Studies in Yue Dialects 1: Phonology of Cantonese*. Cambridge University Press.

Hồ Ngọc Đức. 2004. Vietnamese word list. http://www.informatik.uni-leipzig.de/∼duc/software/misc/wordlist.html. Accessed 24 February 2021.