[In:] Hans Sauer, Piotr P. Chruszczewski (eds.) *Mostly Medieval: In Memory of Jacek Fisiak*.
(Beyond Language 5). San Diego, CA: Æ Academic; 445–458.

THE NAD PHONOTACTIC CALCULATOR: AN ONLINE TOOL TO CALCULATE CLUSTER PREFERABILITY ACROSS LANGUAGES

Katarzyna Dziubalska-Kołaczyk and Dawid Pietrala

ABSTRACT: This paper presents the Net Auditory Distance (NAD) phonotactic calculator which can be used for calculating preferability of consonant clusters. The concept of measuring the preferability is grounded in the NAD model of phonotactics, which derives from the theory of Beats-and-Binding Phonology (Dziubalska-Kolaczyk 2002): NAD calculation is based on the computation of relations between place of articulation (POA), manner of articulation (MOA), and sonorant-obstrucent dichotomy of sounds which are part of the cluster. The values of the parameters are language-specific. So far, the values for the calculations in seven languages have been made available in the calculator: Polish, English, German, Russian, Ukrainian, Croatian and Georgian. The tool is accessible online at http://wa.amu.edu.pl/nadcalc/ and is regularly updated by the authors.

KEYWORDS: phonotactics, phonological preferability, tool, consonant cluster, NAD

1. BACKGROUND

The Beats-and-Binding Phonology (B&B henceforth) is a theory of phonology proposed by Dziubalska-Kołaczyk (1995, 2002a, 2009, *etc.*). B&B is a syllable-less model of phonology which is derived from the framework of Natural Linguistics and Natural Phonology.[1] The B&B phonotactics constitutes a crucial part of the entire model because it accounts for the organization of consonant clusters in language. The model also aims at predicting the behavior of consonant clusters in first and second language acquisition, speech disorders and diachronic change. The B&B model replaces traditional syllable components (onset, nucleus, coda) with beats and non-beats: beats

[1] Stampe (1969, 1979); Donegan & Stampe (1979); Dressler (1985); Dziubalska-Kołaczyk (1995, 2001a, 2001b, 2002a, 2002b, 2002c, 2007) and Dressler & Dziubalska-Kołaczyk (2006).

are best realized by vowels and are relatively more prominent, while non-beats are best realized by consonants and are relatively less prominent.

Vowels typically act as beats as they are more salient thanks to high sonority and articulatory openness, however, consonants may at times be beats as well (for instance, syllabic consonants in English). Beats and non-beats form relations called bindings. Bindings connect beats (B) and non-beats (nB) in a sequence. Typically, sequences of beats and non-beats alternate so as to avoid hiatuses, *i.e.* the contact of two beats. This is often achieved by the insertion of consonants (non-beats) between the two beats.

The syllable in B&B Phonology is epiphenomenal or, at the most, emergent due to principled phonotactic preferences, also referred to as universal and language-specific preferability generalizations. The phonotactic preferences are responsible for different degrees of intersegmental cohesion which, in turn, determines the behavior of segments and creates the impression of syllable structure.

2. MEASURING CLUSTER PREFERABILITY

The suggested measure of cluster preferability is based on phonetic contrast: sounds are more easily recognized, learned, and transmitted when they contrast well with their neighbors. From this perspective, CV sequences are more preferred than the sequences of two or more consonants, *i.e.* clusters; they are learned and transmitted more easily and are historically more stable. Among clusters, those of lower degree of preferability (in other words, more marked) are less stable and more prone to disappear than the more preferred ones.

In order to evaluate the preferability of a cluster, we need to (1) calculate the cluster's NAD (net auditory distance) values for all combinations of consonants in a cluster as well as the sequence of a consonant and its adjacent vowel and (2) assess the NAD values against the phonotactic preference condition.

The NAD value of a sequence of sounds is calculated on the basis of the manner of articulation, place of articulation, and sonorant/obstruent distinction for each of the sounds: MOA + POA + S/O (the S/O distinction replaced the Lx (voiced/voiceless) distinction used in our previous research). Values for MOA and POA can be read from tables provided by Dziubalska-Kołaczyk (2014: 10):

POLISH CONSONANTS

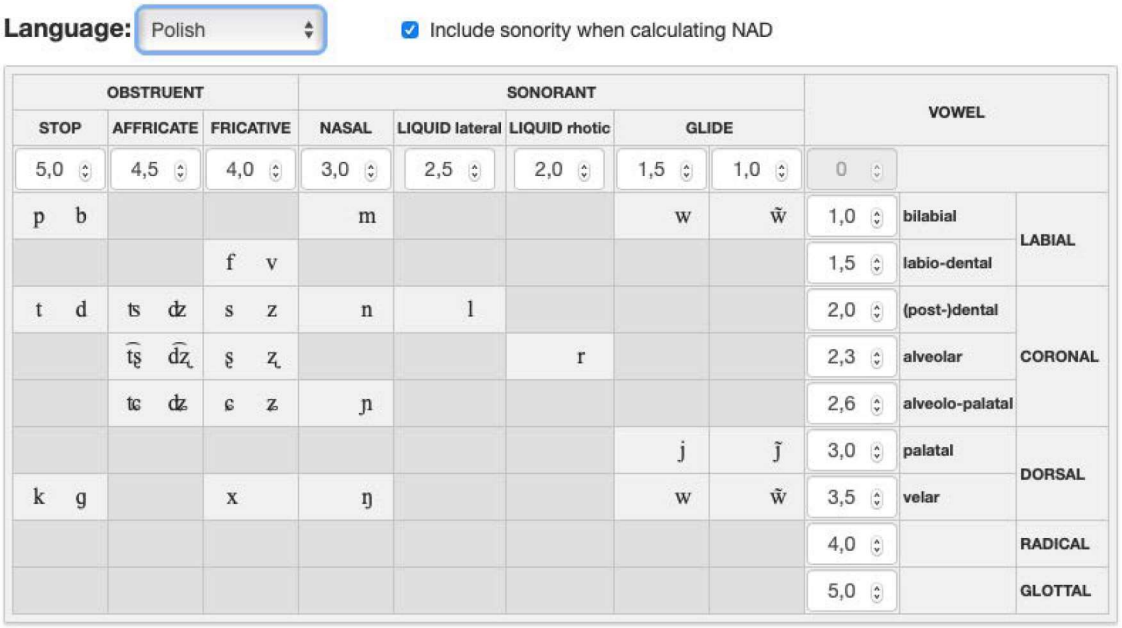

Fig. 1. MOA and POA values for Polish

For instance, for a Polish cluster /prV/ (where 'v' stands for a vowel), we can determine that

MOA for /p/ = 5;
MOA for /r/ = 2;
POA for /p/ = 1; and
POA for /r/ = 2.3.

For vowels, MOA is 0. Then, in order to calculate NAD, the following operations must be performed:

1. /pr/ = |(MOA1 – MOA2| + |(POA1 – POA2)| + S/O
2. /pr/ = |5 – 2| + |1 – 2.3| + 1 = 3 + 1.3 + 1 = 5.3
3. **NAD CC = 5.3**
4. /rV/ = |MOA1 – MOA2|
5. /rV/ = |2 – 0| = 2
6. **NAD CV = 2**

ENGLISH CONSONANTS

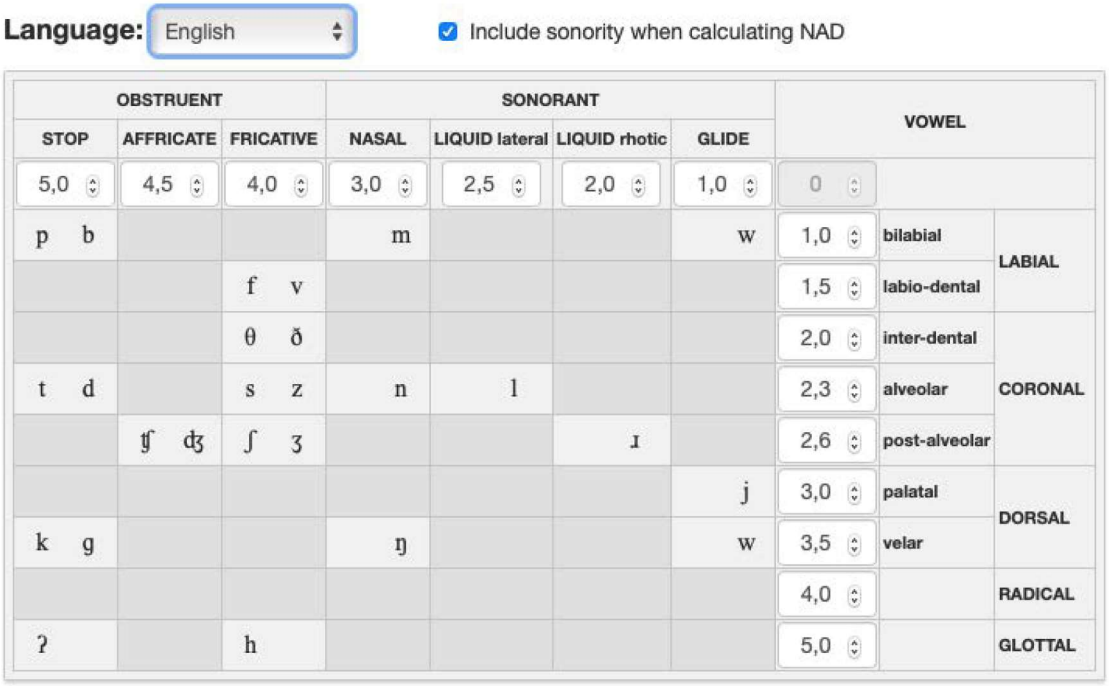

Fig. 2. MOA and POA values for English

Analogically, for an English initial cluster /pɹV/,

MOA: /p/ = 5, /ɹ/ = 2;
POA: /p/ = 1, /ɹ/ = 2.6.

The calculation for the cluster looks as follows:

1. /pɹ/ = $|5 - 2| + |1 - 2.6| + 1 = 3 + 1.6 + 1 = 5.6$ (NAD: 5.6)
2. /ɹV/ = $|2 - 0| = 2$ (NAD: 2)

Knowing the NAD values, we can check if a double initial cluster is preferred on the basis of the following condition:

C1C2V	NAD (C1,C2) ≥ NAD (C2,V)

Fig. 3. The phonotactic preference condition for double initial clusters

The condition reads:

> In word-initial double clusters, the net auditory distance (NAD) between the two consonants should be greater than or equal to the net auditory distance between a vowel and a consonant neighbouring on it.

Thus, for our two examples, we can conclude that:

1. /pr/ in Polish is regarded as preferred as it conforms the condition NAD CC > NAD CV (5.3 > 2);
2. /pɹ/ in English is regarded as preferred as it also conforms the above condition (5.6 > 2).

The NAD conditions for the remaining word positions are specified by the well-formedness conditions listed below. The condition for double final clusters states:

VC1C2	NAD (V,C1) ≤ NAD (C1,C2)

Fig. 4. The phonotactic preference condition for double final clusters

The condition reads:

> In word-final double clusters the net auditory distance (NAD) between the two consonants should be greater than or equal to the sonority distance between a vowel and a consonant neighbouring on it.

The condition for double medial clusters states:

V1C1C2V2	NAD (V1, C1) ≥ NAD (C1,C2) ≤ NAD (C2,V2)

Fig. 5. The phonotactic preference condition for double medial clusters

The condition reads:

> For a word-medial double cluster, the NAD between the two consonants should be less than between each of the consonants and its respective neighbouring beat, and it may be equal to the NAD between the first consonant and the beat preceding it.

The predictions for triple clusters in all word positions are formulated below. The condition for triple initials states that:

C1C2C3V	NAD (C1,C2) < NAD (C2,C3) ≥ NAD (C3,V)

Fig. 6. The phonotactic preference condition for triple initial clusters

The condition reads:

> For word-initial triple clusters, the NAD between the third consonant and the second consonant should be greater than or equal to the NAD between

this third consonant and the vowel, and greater than the NAD between the second and the first consonant.

The condition for triple finals states that:

VC1C2C3	NAD (V,C1) ≤ NAD (C1,C2) > NAD (C2, C3)

Fig. 7. The phonotactic preference condition for triple final clusters

The condition reads:

For word-final triple clusters, the NAD between the first consonant and the second consonant should be greater than or equal to the NAD between this first consonant and the beat, and greater than the NAD between the second and the third consonant.

The condition for triple medials states that:

VC1C2C3V	NAD (V,C1) ≥ NAD (C1,C2) & NAD (C2, C3) < NAD (C3, V2)

Fig. 8. The phonotactic preference condition for triple medial clusters

The condition reads:

For word-medial triple clusters, the NAD between the first and the second consonant should be less than or equal to the NAD between the first consonant and the beat to which it is bound, whereas the NAD between the second and the third consonant should be less than between the third consonant and the beat to which it is bound.

It can be said that the more a cluster diverges from the preference, the more marked (less preferred) it gets.

3. NAD REVISED

The values for MOA and POA described above are based on the Ladefoged's MOA and POA classifications (Ladefoged 2014), which were assigned numerical values. These classifications have been adjusted to account for the fact that not all distinctions are equally strong in the phonologies of all languages and some languages may make finer phonological distinctions than others (Dziubalska-Kołaczyk 2019).

However, the NAD preferability measure is not tied to these particular values. Right now, research is being conducted to use more fine-grained distinctions. An example of this is the use of the hierarchy of relative sonority proposed by Parker (2008: 60). The set of the MOA and POA values in this approach would be:

POLISH CONSONANTS

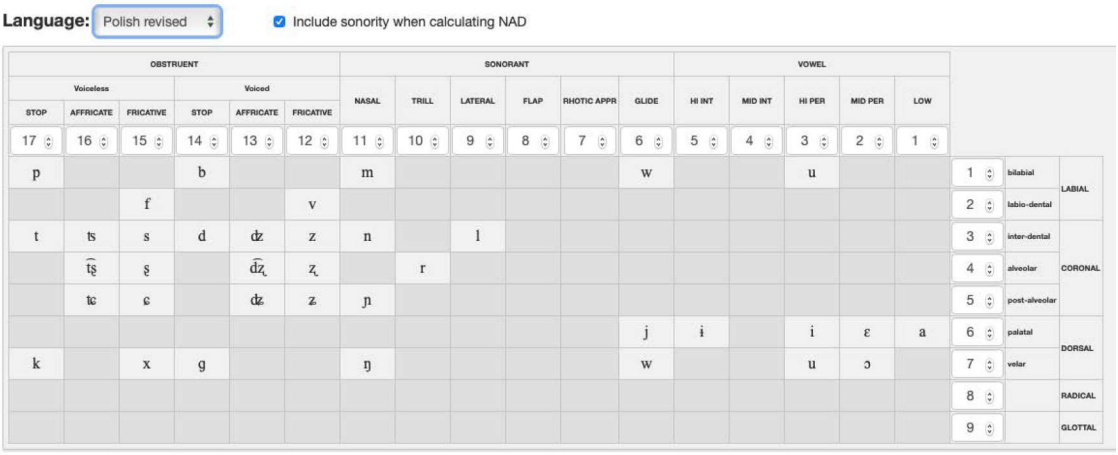

Fig. 9. Revised MOA and POA for Polish

ENGLISH CONSONANTS

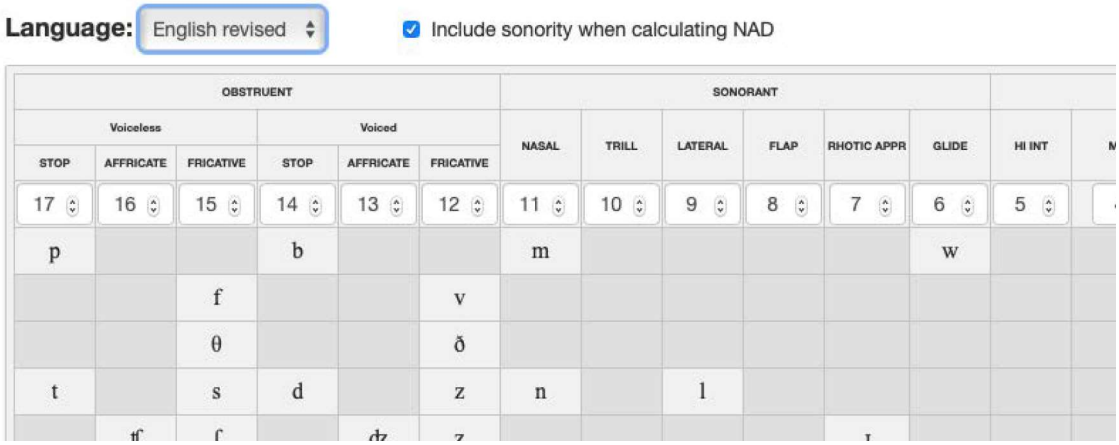

Fig. 10. Revised MOA and POA for English

This set of values is being validated to determine its appropriacy for NAD and cluster preferability measurements.

4. ONLINE TOOL FOR CALCULATING CLUSTER PREFERABILITY

As could be seen above, calculating the NAD between two segments involves a series of operations which are purely arithmetic in nature. To aid automate calculations several tools have been developed which process input provided by the user and return NAD values along with preferability ratings. The first version of the NAD Phonotactic Calculator (1.0) was developed by Grzegorz Krynicki in the years 2006–2007 (Dziubalska-Kołaczyk & Krynicki 2007). It was an online application written in the Perl language. The following version of the tool developed by Dawid Pietrala was largely based on version 1.0+ in terms of functionality and interface design but it was transformed into a desktop application with the option to modify the MOA and POA values for sounds (Pietrala 2014). The most recent version of the tool (3.0), built by Grzegorz Aperliński (*cf.* 2019), was made an online tool again. It is publicly available at http://wa.amu.edu.pl/nadcalc/. This version combines the features of the two previous versions: it is an online application which also provides options to modify phone values in the online environment. In addition, it introduces some additional functions such as being able to calculate NAD for several clusters at a time as well as being able to calculate NAD for languages other than English and Polish. It also offers a cleaner and simpler interface.

5. HOW TO USE THE ONLINE TOOL

The tool can be used by researchers who work in the framework of B&B phonotactics to perform calculations of NAD in a robust way. To compute the preferability of a given cluster, the user simply writes it in the input box. The user does not have to take care about the values of POA or MOA as the tool is distributed with a database of default values for specific phonemes as well as with a set of methods which perform the necessary mathematical operations. The application automatically carries out all the calculations and provides the user with the NAD values along with a visual indication of the preferability of the cluster (green or red).

The input can be specified by the user either by writing or copying and pasting the text into the input box, or by clicking the IPA symbols located in the table next to the input box. It is possible to calculate the preferability of more than one cluster at a time; this is done by writing one cluster per line. The tool makes use of standard IPA symbols to identify individual consonants

in a cluster. For the original B&B model, vowels are identified with the use of the V symbol. This was because previously the B&B model did not distinguish between vowels when measuring cluster preferability. The revised NAD, however, does assign specific values which depend on the vowel positions and the values are demonstrated above (*cf.* Figures 9 and 10). Providing the V symbol in a cluster is necessary in order to specify the position of (a) vowel(s) relative to the cluster and, consequently, the word position of the cluster. For example, the string stV would be recognized as a word-initial /st/ cluster, while Vst would be treated as a word-final /st/ cluster. The tool also recognizes word-medial clusters, *e.g.* VstV. Both two- and three-segment clusters are supported in all word positions. The program validates user input so that it returns a user-friendly message when incompatible input is given.

The user can define a number of settings when using the tool. A drop-down list located above the IPA symbols provides a list of supported languages. At the time of writing this paper, the program supports Polish, English, Polish revised (based on Parker 2008), English revised (based on Parker 2008), German, Russian, Ukrainian, Georgian, Croatian, as well as generic user-defined set of parameter values. The default language at this point is English (with values based on Ladefoged 2014). When the user chooses a language, the program displays the possible segments for that language along with the default POA and MOA values for the segments. The user can make temporary changes to these values if necessary. The modifications of POA and MOA values for specific languages are saved for the duration of the current session and revert to default when reloading the webpage. The modifications made when using the user-defined option are saved in the web-browser and will be retrieved when accessing the application in the future. It is also possible to control the way NAD is calculated by deselecting the checkbox next to the language selection list. By turning this option off, the program will not take into account sonority when computing the preferability of clusters and will estimate NAD on the basis of MOA and POA.

The program operates based on the input provided by the user and the current values of POA and MOA parameters. When the user clicks the calculate button underneath the input box, the program will automatically carry out the necessary calculations. First, the program checks whether the segments provided are allowed for the specified language. Then, the CV structure of the cluster is identified based on the number of segments and the position of the vowel(s). Next, the NAD of each pair of segments is calculated on the basis of the currently specified parameter values. Finally,

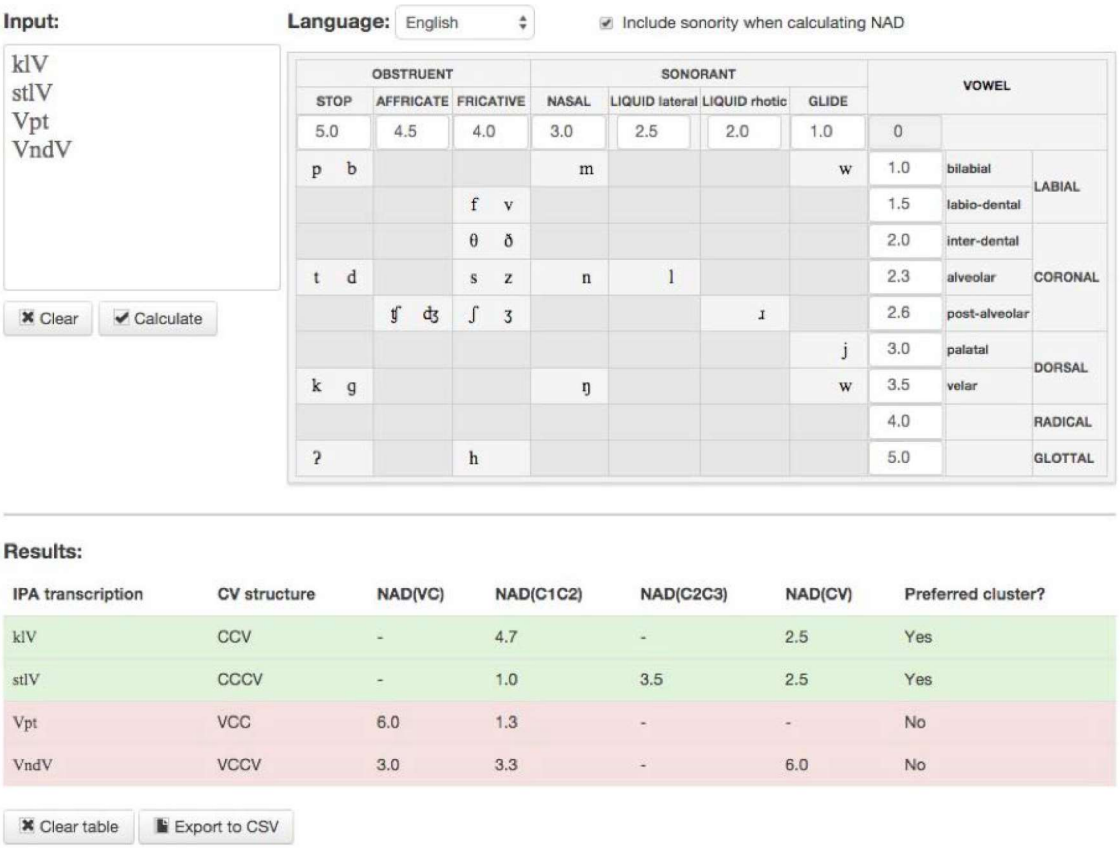

Fig. 11. Phonotactic preference calculator along with sample calculations

cluster preferability is measured with the help of the corresponding formula. The results are displayed in a table at the bottom of the screen and can be exported to a .csv file.

6. CONSIDERATIONS TO KEEP IN MIND

The purpose of the application is to aid researchers in calculating cluster preferability and experimenting with different parameter values for different languages. By offering automatic computation of cluster preferability, the program does away with the laborious task of measuring NAD by hand and limits the possibility of human error in calculations. However, there are some considerations that need to be taken into account when using this tool. The most important consideration is related to one of the reasons the new version of the program has been developed: to facilitate computations of cluster preferability for languages other than English and Polish. The user can set their own parameter values for segments present in the language they are

researching by choosing the user-defined language option. To keep value manipulations simple, the user is presented with a chart containing the majority of standard IPA symbols. By default, parameter values are empty but can be modified by the user by specifying a number between 0.0 and 5.0 for POA and MOA. Thanks to this, only the values for segments that are relevant for a given language need to be specified. The user-friendliness of this solution, however, comes at a price. First of all, the position of all segments in the chart is fixed and cannot be changed by the user. The positions of segments have been based on places and manners of articulation (Selkirk 1984; Ladefoged 2014). Nonetheless, the order of the segments does not reflect the hierarchy and organization of segments in all languages. Another limitation imposed by the design is that it is not possible to add new segments that are not covered by the existing chart. Because of the above, the online NAD calculator is not suited to be used with all languages of the world. Users are thus encouraged to contact the authors with comments relating to the layout of the chart as well as with suggested POA and MOA values for new languages.


Please note that the position of the consonants in the table below is purely **functional**.
If you have any comments or suggestions about the grouping of the consonants, please direct them to the authors.

OBSTRUENT			SONORANT						VOWEL		
STOP	AFFRICATE	FRICATIVE	NASAL	LIQUID lateral	LIQUID rhotic	TAP or FLAP	TRILL	GLIDE			
p b		ɸ β	m				ʙ	ʌ̟ w		bilabial	LABIAL
		f v	ɱ			ⱱ		ʋ		labio-dental	
t̪ d̪	t̪s̪ d̪z̪	θ ð								dental	CORONAL
t d	ʦ ʣ	s z	n	l	ɹ	ɾ	r			alveolar	
	ʧ ʤ	ʃ ʒ			ɹ					post-alveolar	
ʈ ɖ		ʂ ʐ	ɳ	ɭ	ɻ	ʈ				retroflex	
				ʎ				j		palatal	DORSAL
k g			ŋ	ʟ				ʌ̟ w		velar	
										uvular	RADICAL
										epiglottal	
ʔ		h								glottal	GLOTTAL

Fig. 12. Custom IPA chart with spaces for user-defined parameter values

Another feature of the software that needs to be considered is its session retrieval capability. All changes made in the user-defined IPA chart are automatically saved in the local storage of the currently used internet browser.

Thanks to this solution, user validation is not required to retrieve modifications made to the parameter values. However, these changes will not be carried over when using a different browser or different machine.

One final design choice to mention is the organization of the output produced by the software. Version 3.0 provides the NAD values for all consonant pairs in the cluster as well as for the first and/or last consonant and the adjacent vowel. Previous version also provided the results of the calculations used to assess cluster preferability. These results could be used to evaluate the relative preferability of two-consonant clusters: the higher the value the more preferable the cluster. However, in the case of three-consonant clusters such evaluations were not possible. To maintain consistency of the results for all cluster types, these calculations are not provided in the output and instead a more general cluster preferability rating is displayed (either 'yes' or 'no') along with color-coding (green for preferable and red for not preferable). More detailed calculations can be carried out on exported .csv files in software such as Microsoft Excel.

ABBREVIATIONS

B&B – Beats-and-Binding
MOA – manner of articulation
NAD – Net Auditory Distance
POA – place of articulation

REFERENCES

Aperliński, Grzegorz (2019) NAD Phonotactic Calculator 3.0. [online tool http://wa.amu.edu.pl/nadcalc/; date of access: 11th March 2020]

Donegan, Patricia Jane, David **Stampe** (1979) "The Study of Natural Phonology." [In:] Daniel A. Dinnsen (ed.) *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press; 126–173. [Also digitalized at: https://publish.iupress.indiana.edu/read/current-approaches-to-phonogical-theory/section/47173ac2-a5f8-47a9-98a7-b9f7e50664f7]

Dressler, Wolfgang U (1985) *Morphonology: The Dynamics of Derivation*. Ann Arbor, MI: Karoma Press.

Dressler, Wolfgang U., Katarzyna **Dziubalska-Kołaczyk** (2006) "Proposing Morphonotactics." [In:] *Italian Journal of Linguistics (IJL) | Rivista di Linguistica 18*(2); 249–266.

Dziubalska-Kołaczyk, Katarzyna (1995) *Phonology Without the Syllable: A Study in the Natural Framework*. Poznań: Motivex.

Dziubalska-Kołaczyk, Katarzyna (2001a) "Phonotactic Constraints are Preferences." [In:] Katarzyna Dziubalska-Kołaczyk (ed.) *Constraints and Preferences*. Berlin, Germany | Boston, MA: De Gruyter Mouton; 69–100. DOI: 10.1515/9783110881066

Dziubalska-Kołaczyk, Katarzyna (2001b) "On the Phonotactics of Consonantal Clusters." [In:] Chris Schaner-Wolles, John Rennison, Friedrich Neubarth (eds.) *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*. Torino: Rosenberg & Sellier, 111–119.

Dziubalska-Kołaczyk, Katarzyna (2002a) *Beats-and-Binding Phonology*. Bern, Switzerland: Peter Lang.

Dziubalska-Kołaczyk, Katarzyna (2002b) "Challenges for Natural Linguistics in the Twenty-first Century: A Personal View." [In:] *University of Hawai`i Working Papers in Linguistics 23*; 15–39.

Dziubalska-Kołaczyk, Katarzyna (2002c) "Challenges for Natural Linguistics in the Twenty-first Century: A Personal View." [In:] Katarzyna Dziubalska-Kołaczyk, Jarosław Weckwerth (eds.) *Future Challenges for Natural Linguistics*. Munich, Germany: Lincom Europa; 103–128.

Dziubalska-Kołaczyk, Katarzyna (2007) "Natural Phonology: Universal Principles for the Study of Language: Insiders Meet Outsiders." [In:] Jürgen Trouvain, William J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS), Saarbrücken, 6–10 August 2007*. Saarbruecken, Germany: University of Saarland in Saarbruecken [Saarbrücken Univ. des Saarlandes], 71–75.

Dziubalska-Kołaczyk, Katarzyna (2009) "NP Extension: B&B Phonotactics." [In:] *Poznań Studies in Contemporary Linguistics 45*(1); 55–71. DOI: 10.2478/v10010-009-0011-9

Dziubalska-Kołaczyk, Katarzyna (2014) "Explaining Phonotactics using NAD." [In:] *Language Sciences 46*(A); 6–17. DOI: 10.1016/j.langsci.2014.06.003

Dziubalska-Kołaczyk, Katarzyna (2019) "On the Structure, Survival and Change of Consonant Clusters." [In:] *Folia Linguistica Historica 40*(1); 107–127. DOI: 10.1515/flih-2019-0006

Dziubalska-Kołaczyk, Katarzyna, Grzegorz **Krynicki** (2007) "Predicting Difficulty in Second Language Acquisition." [A paper presented at *New Sounds 2007*, November 25–28, 2007].

Ladefoged, Peter [Keith Johnson] ([1975] 2014) *A Course in Phonetics*. 7th ed. Belmont, CA: Cengage Learning.

Parker, Steve (2008) "Sound Level Protrusions as Physical Correlates of Sonority." [In:] *Journal of Phonetics 36*(1); 55–90. DOI: 10.1016/j.wocn.2007.09.003

Pietrala, Dawid (2014) *The Application of Linguistic Principles in Programming: Phonemic Transcription and Phonotactic Constraints* (Unpublished PhD dissertation). Poznań, Poland: Adam Mickiewicz University in Poznań.

Selkirk, Elisabeth (1984) "On the Major Class Features and Syllable Theory." [In:] Mark Aronoff, Richard T. Oehrle (eds.) *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students*. Cambridge, MA: MIT Press; 107–113.

Stampe, David (1969) "The Acquisition of Phonetic Representation." [In:] Robert Binnick, Alice Davidson, Georgia M. Green, Jerry L. Morgan (eds.) *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*. Chicago, IL: University of Chicago; 443–453. [Reprinted in: Barbara Lust, Claire Foley (eds.) (2004) *First Language Acquisition: The Essential Readings*. Chicago, IL: Chicago Linguistic Society; 307–316.]

Stampe, David (1979) *A Dissertation on Natural Phonology*. (Outstanding Dissertations in Linguistics 22). New York, NY: Garland Publishing.