Information theoretic syllable structure and its relation to the c-center effect

Uwe D. Reichel

Institute of Phonetics and Speech Processing, University of Munich

reichelu@phonetik.uni-muenchen.de

Abstract

Established phonological theories postulate uniform syllable constituent structures. From a traditional hierarchical point of view, syllables are right branching implying a close connection between the nucleus and the coda. Articulatory Phonology in contrast suggests a stronger cohesion between onsets and nuclei than between nuclei and codas. This claim is empirically supported by the c-center effect which initially has been observed for onsets only. Nevertheless, recent studies revealed that this effect does not occur in all complex onsets and can also be observed in codas. To account for this structure non-uniformity, we propose an information theoretic approach to measure connection strengths between syllable constituents in terms of their pointwise mutual information. It turned out that the derived constituent structures correspond well to the empirical c-center findings on American English and German data. The results are discussed from a Usage-based Phonology perspective considering c-centers to be a frequency effect.

Index Terms: syllable structure, c-center, information theory, Articulatory Phonology, Usage-based Phonology, frequency effect

10.21437/Interspeech.2013-96

1. Introduction

From a traditional hierarchical (TH) point of view the syllable can be divided into two constituents, the onset and the rhyme the latter dominating the nucleus and the coda [1, 2] (cf. Figure 1 A) . The underlying assumption is that nucleus and coda are more strongly connected than onset and nucleus. The TH view is mainly supported by the observation that nucleus and coda together are relevant for rhyming judgments, and that the syllable weight constraining word stress assignment in many languages (see e.g. [3] for German) is only determined by nucleus and coda, but not by the onset.

In contrast to TH, Articulatory Phonology (AP) claims a stronger binding between onsets and nuclei than between nuclei and codas [4] (cf. Figure 1 B). These bindings are expressed in form of a uniform gestural coupling pattern within syllables as follows: all onset consonants but only the first coda consonant are coupled to the vowel. Empirical evidence is given by the c-center effect, that is generally observed for syllable onsets but not for codas [4, 5, 6]. The temporal distance between the c-center (the time midpoint) of the syllable onset to the nucleus has been observed to be constant for different onset cluster lengths, implying that the gestural overlap between onset consonants increases with increasing segment number. This overlap results from a compromise among the competing onset consonants so that each consonant can retain its in-phase coupling to the vowel as far as possible. Coda segments in contrast generally do not tend to overlap more with increasing consonant number, since only the first segment is coupled to the vowel. Thus they do not pertain a constant c-center.

Note, that the term *c-center effect* in the following is used without any implications on phasing of the gestural coupling but only to describe the surface phenomenon that a constant time interval between nucleus and cluster midpoint is pertained.

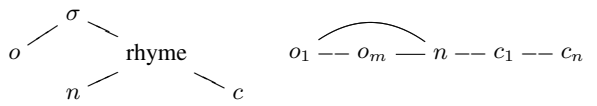

Figure 1: **A (left):** Hierarchical right-branching syllable constituent structure of onset o, nucleus n and coda c resulting in a stronger connection between nucleus and coda ($o|nc$ pattern). **B (right):** Coupling between syllable constituents $o = o_1 \dots o_m$, n, and $c = c_1 \dots c_n$ according to AP (solid: in-phase, dashed: anti-phase). The onset is entirely coupled and thus closer connected to the nucleus than the coda ($on|c$ pattern).

The outcome of recent studies shed doubt on a uniform AP handling of constituent connection strength across all syllable types. For American English Marin and Pouplier [7] observed as expected a c-center effect for the onset clusters $/kl\ pl\ sk\ sm\ sp/$, but also in contradiction to AP predictions for the coda clusters $/lk\ lp/$ (in its temporal surface related meaning the term is used here, i.e. without phasing implications).

For German Pouplier [8] found as expected a c-center effect for the onsets $/bl\ sk/$, but against the AP predictions not for the onset $/pl/$. These studies show, that the c-center effect (1) can be observed not only in onsets but also in codas, and (2) it does not occur in all onsets.

Thus, the connection strength between syllable constituents cannot be addressed uniformly over all syllable types. To allow for a more flexible access we assume a flat syllable structure based on the word games findings of [9] and propose an information theoretic (IT) approach to quantify the constituents' connection strengths in dependency of the segmental content of the syllables. Two distinct constituent patterns will be distinguished: $on|c$ represents a stronger connection between onset o and nucleus n as postulated by AP. $o|nc$ stands for a stronger connection of nucleus and coda as postulated by TH.

In section 2 we introduce our approach in greater detail. Its application on the findings of [7] and [8] as well as the results are described in sections 3 and 4. In section 5 the results will be discussed from a usage-based perspective.

2. Information theoretic syllable structure induction

2.1. General constituent pattern trend

To measure the general connection strength of syllable constituents in a language we employed the mutual information

measure. The symmetric mutual information $I(X,Y)$ between X and Y gives the amount of uncertainty reduction to predict the value of the variable X if the value of Y is given and vice versa. Adopted to syllable constituents

$$\begin{aligned} I(N,O) &= H(N) - H(N|O) \\ I(N,C) &= H(N) - H(N|C) \end{aligned}$$

give the overall mutual information between the set of nucleus types N and onset types O of a language, and between N and coda types C. $H(N)$ stands for the entropy of N, and $H(N|O)$ for its conditional entropy given the onset is known.

If $I(N,O) > I(N,C)$, the co-occurrence of onsets and nuclei is more regular than between nuclei and codas, thus from an information theoretic point of view the general constituent pattern trend is $ON|C$, i.e. the connection strength between onsets and nuclei is generally higher than between nuclei and codas. For $I(N,O) < I(N,C)$ the trend is $O|NC$.

2.2. Single constituent patterns

The pointwise mutual information PMI quantifies the connection strength between particular syllable constituents n, c, o:

$$\begin{aligned} PMI(n,o) &= \log_2 \frac{p(n,o)}{p(n)p(o)} \\ PMI(n,c) &= \log_2 \frac{p(n,c)}{p(n)p(c)} \end{aligned}$$

Analogously to the preceeding section, in case of $PMI(n,c) > PMI(n,o)$, the IT predicted pattern is $on|c$, while $PMI(n,c) < PMI(n,o)$ yields the pattern $o|nc$. Since PMI overestimates the coherence of low-frequency events (due to small denominator values) a frequency threshold has to be set for o, n, and c.

2.3. Decision tree induction

In order to get classifiers for a syllable's constituent pattern from its constituents, we built information gain decision trees ([10], adapted from [11]). The input material comprised feature vectors containing syllable onset o, nucleus n and coda c, together with the associated target value ($on|c$ or $o|nc$) which is derived from comparing $PMI(n,c)$ and $PMI(n,o)$. As an example, the syllable [bVlk] is represented as $< [b, V, lk], o|nc >$ since $PMI(/b/, /V/) = 1.55 < PMI(/V/, /lk/) = 3.13$. The trees are created by recursive partitioning of syllables with respect to the constituent which contributes the highest information gain about the target pattern. As a result, the trees assign to each observed syllable type the associated pattern. Two examples are given in Figures 2 and 3.

3. Model application

We applied our approach to the syllable inventory of English (referred to as *eng* in the following) and German (*deu*) in general as well as more specifically to the studies of [7] and [8]. Only syllables containing both onset and coda were considered. For American English, [7] found a c-center effect for the onsets $/kl\ pl\ sk\ sm\ sp/$ but also for the codas $/lk\ lp/$. For German, [8] reported a c-center effect for the onsets $/bl\ sk/$ but not for the onset $/pl/$. Thus both studies' outcomes partly contradict the AP predictions. Table 1 summarizes these findings on the syllable sets in the following referred to as eng_1, eng_2, deu_1, and deu_2, respectively. The transcriptions are given in UK and German SAMPA.

Table 1: *Schematic summarization of empirical c-center findings on the connection strength of onset o, nucleus n and coda c.* CP*: syllable constituent pattern, | separates the constituents which are less strongly connected.*

set	**syllable types**	***CP***
eng_1	$o = /kl\ pl\ sk\ sm\ sp/$	$on\|c$
eng_2	$c = /lk\ lp/$	$o\|nc$
deu_1	$o = /bl\ sk/$	$on\|c$
deu_2	$o = /pl/$	$o\|nc$

[8] additionally examined the onsets $/gm\ km/$ which were discarded in the current study since they do not allow for reliable PMI calculations due to their low frequencies. Slight SAMPA differences between the American and UK variant are considered to be negligible for our study and do not affect the transcriptions of the stimulus words used in [7].

3.1. Data

The overall mutual information between all onsets, nuclei and codas as well as all pointwise mutual informations between onset, nucleus, and coda types were calculated on the CELEX lemma pronunciation dictionaries [12] for English and German, respectively. To allow for both $on|c$ and $o|nc$ patterns, only those syllables containing both onset and coda were considered. The syllables were split into their constituents and the constituent count increments were multiplied by the CELEX-provided frequencies of the words they belong to. This weighting serves to receive counts related to word tokens instead of types and thus more usage-related frequency distributions. By this, the training data comprised 56905 syllable tokens for English and 77625 tokens for German (cf. table 2). Syllable constituents below a frequency threshold of 10 were discarded.

Table 2: *Syllable token and constituent type counts. The type counts are determined by the syllable set definitions in table 1, e.g. in set* eng_2 *the 2 coda types* $c = /lk\ lp/$ *co-occur with 12 onset and 4 nucleus types.*

set	**syllable tokens**	**constituent types**		
		onsets	**nuclei**	**codas**
eng	56905	93	23	118
eng_1	1732	5	19	55
eng_2	76	12	4	2
deu	77625	85	37	96
deu_1	403	2	16	19
deu_2	244	1	12	19

3.2. Structure induction

We built decision trees one for each data set $eng_{1,2}$ and $deu_{1,2}$ mapping the syllable types to the constituency patterns $on|c$ or $o|nc$. The trees were not pruned so that they formed a compact representation of the observed syllable data together with the PMI induced constituent pattern. Tree examples for the syllable sets eng_2 and deu_1 are shown in the Figures 2 and 3.

Based on the tree outputs we calculated the probabilities to belong to an $on|c$ or an $o|nc$ syllable for each onset and coda type.

```
o = /b f g h j k m p s t w/: o|nc
o = /sk/
|  n = /V/: o|nc
|  n = /{/: on|c
```

Figure 2: Decision tree for constituent pattern prediction for English codas */lk lp/* [7]. To be read from left to right. E.g. [bVlk] $\longrightarrow$ *o|nc* (row 1).

```
n = /2: O u:/: on|c
n = /E: a: aI e: Y i:6/: o|nc
n = /E/
|  c = /f l n nt p s t tS x/: on|c
|  c = /k/: o|nc
n = /I/
|  c = /N Nk k ks n nt s ts/: on|c
|  c = /S x/: o|nc
n = /U/
|  c = /lp/: on|c
|  c = /N Nk f/: o|nc
n = /a/
|  c = /N l n s st t/: on|c
|  c = /Nk f lp m/: o|nc
n = /i:/
|  c = /m/: on|c
|  c = /k/: o|nc
n = /o:/
|  c = /p/: on|c
|  c = /s/: o|nc
n = /y:/
|  c = /s/: on|c
|  c = /m/: o|nc
```

Figure 3: Decision tree for constituent pattern prediction for German onsets */bl sk/* [8]. E.g. [blEn] $\longrightarrow$ *on|c* (row 4).

4. Results

4.1. General constituent pattern trend

The general trend based on mutual information is shown in table 3. It can be seen that for the complete syllable inventory the IT approach follows the claims of TH and not of AP. This is also reflected in the PMI distributions for all syllable types in Figure 4 again showing an overall *o|nc* tendency for both languages.

Table 3: *General constituent pattern* CP *trend for English and German based on Mutual information.*

language	I(N,O)	I(N,C)	*CP*
eng	0.41	0.71	*o|nc*
deu	0.46	0.87	*o|nc*

4.2. Selected onset and coda types

For the onset and coda types of the two selected studies the PMI value distributions as well as the probability distributions of the types to be part of an *on|c* or an *o|nc* syllable are presented in Figures 5 and 6. All relations between the examined constituents and the *on|c* vs. *o|nc* pattern assignment are highly significant ($47 < \chi^2_1 < 317, p = 0.001$).

It turned out, that:

- In English syllables with the onsets */kl pl sk sm sp/* (*eng*$_1$) the PMI between onsets and nuclei is higher than between nuclei and codas. This is also reflected in higher probabilities of an *on|c* pattern assignment by the decision tree.
- In English syllables with the codas */lk lp/* (*eng*$_2$) the PMI between nuclei and codas exceeds the PMI between onsets and nuclei, reflected in higher *o|nc* pattern assignment probabilities by the tree.
- In German */bl sk/* onset syllables (*deu*$_1$), onsets are more strongly connected to nuclei in terms of PMI than codas resulting in higher *on|c* assignment probabilities.
- Also gradual overlap differences are reflected in the output probabilities of the *deu*$_1$ tree in Figure 3. */sk/* onsets are reported to be more overlapped than */bl/* onsets by [8] which corresponds well to the tree predictions. 89% of the */sk/* onset syllable tokens but only 82% of */bl/* onset syllables are classified as *on|c*.
- In contrast, in German */pl/* onset syllables (*deu*$_2$), onset-nucleus PMIs are lower than nucleus-coda PMIs leading to a higher *o|nc* assignment probability.

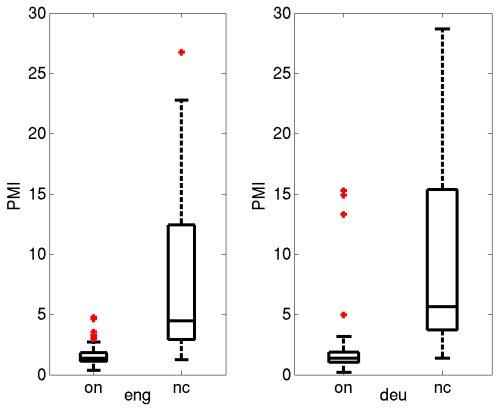

Figure 4: *Pointwise mutual information PMI for all onset-nucleus (o, n) and nucleus-coda (n, c) combinations for English and German.*

In summary, all PMI related constituency pattern probabilities match the findings on the presence and absence of the c-center effect in [7] and [8]. Thus our IT approach provides a better fit to the data than AP and TH as is shown in table 4.

Table 4: *Match between the observed constituent patterns and the predictions of our approach IT, of Articulatory Phonology AP, and of the traditional hierarchical viewpoint TH.*

set	observed pattern	IT	AP	TH
eng$_1$	*on|c* for $o = /kl\ pl\ sk\ sm\ sp/$	+	+	–
eng$_2$	*o|nc* for $c = /lk\ lp/$	+	–	+
deu$_1$	*on|c* for $o = /bl\ sk/$	+	+	–
deu$_2$	*o|nc* for $o = /pl/$	+	–	+

5. Discussion and Conclusion

5.1. Required generalizations

It was possible to predict the presence and absence of a c-center in syllable onsets as well as in codas for two studies on Ameri-

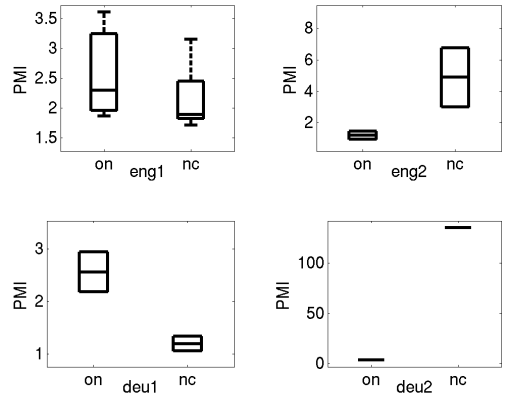

Figure 5: *Pointwise mutual information PMI between onset-nucleus (o, n) and nucleus-coda (n, c) combinations for the syllable sets described in table 1. For each onset type (*eng_1, deu_1, deu_2*), resp. coda type (*eng_2*) the mean* $PMI(o,n)$ *and* $PMI(n,c)$ *were calculated over all syllables this type appeared in. The boxplots represent the distributions of these type-related mean values.*

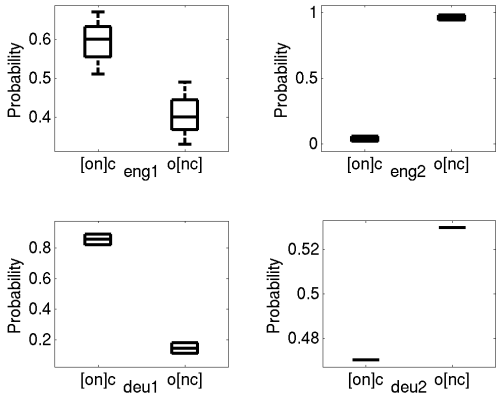

Figure 6: *Probabilities for* $on|c$ *and* $o|nc$ *constituent patterns for the syllable sets described in table 1. For each onset type (*eng_1, deu_1, deu_2*), resp. coda type (*eng_2*) the mean assignment probabilities* $P(\text{"}on|c\text{"})$ *and* $P(\text{"}o|nc\text{"})$ *were calculated over all syllables this type appeared in. The boxplots represent the distributions of these type-related mean values.*

can English and German data. C-centers appear within the constituent that is more closely connected to the nucleus in terms of pointwise mutual information. The set deu_1 requires a generalization of the results to syllable types without onset or coda, since this set contains also one open syllable */blaU/ (blue)*. We propose the following generalization which is in line with the empirical findings: A constituent type that is tightly connected to the nucleus in *onc* syllables retains a c-center in case of its exclusive occurrence in *on* or *nc* syllables, respectively.

The present results are based on constituent counts in *onc* syllables only. Nevertheless, it turned out that an extension of the training to all syllable types did not lead to qualitatively different results for any of the subsets eng_1, eng_2, deu_1, and deu_2. Further generalizations are discussed in the final paragraph.

5.2. Relations to Usage-based Phonology

[13] define the syllable structure as "a characteristic pattern of coordination among gestures". The information theoretic approach itself offers no direct articulatory explanation for this coordination but could provide potential influencing factors from a usage-based viewpoint. Usage-based Phonology (UP) [14] aims to infer phonological processes and categories directly from language use. A crucial UP concept is the *frequency effect* that can trigger elisions and assimilations, morpho-phonological regularizations, phonotactic acceptability judgments, and the organizational integration of frequently co-occurring units like articulatory gestures (examples are given in [14] and [15]). For syllables such frequency effects have been reported and modeled in an exemplar-theoretic framework [16] amongst others by [17] and [18]. They found, that a syllable's duration variability can be inferred from the duration variability of its parts only for low-frequency but not for high-frequency syllables. This discrepancy for the latter is explained by the high co-occurrence frequencies of their segments so that these segments are not anymore accessed individually but are integrated to a single unit in speech production.

Turning back to our study, also the c-center phenomenon could be seen as the result of a frequency effect this time quantified by means of PMI. From this perspective, a high co-occurrence number of a syllable nucleus with an onset or a coda consonant cluster can establish a more integrative gestural organization of this cluster and the nucleus. This integrative organization is expressed in the gestural coupling of *all* consonants of the respective cluster to the nucleus vowel, regardless of whether the cluster forms the syllable onset or the coda. Therefore and in accordance with the reviewed studies the c-center effect is not restricted to the onset anymore, but can be observed in the onset as well as in the coda as an outcome of the hypothesized frequency effect.

It is argued, that opposed to raw co-occurrence counts PMI is a more robust and generalizable measure to quantify frequency effects since it is less dependent on the size of the actual data set and facilitates comparisons of data sets of different sizes.

5.3. Future work

In this initial work our IT approach has been evaluated on four syllable sets only, focusing on studies that reported the violation of Articulatory Phonology predictions of c-centers. Thus, to get a more solid basis, the predictive power of our approach needs confirmation on further data sets – all the more that the reviewed data had not been designed to examine simultaneous c-center effects in onsets and codas, which would principally be possible from the usage-based viewpoint described above. To capture the joint occurrence of two c-centers in one syllable or their complete absence, the current forced-choice decision based on PMI inequality would have to be supplemented or replaced by PMI thresholds to be exceeded to trigger the c-center effect.

6. Acknowledgments

The work of the author has been carried out within the CLARIN-D project [19] (BMBF-funded).

7. References

[1] K. Pike and E. Pike, "Immediate constituents of Mazateco syllables," *International Journal of American Linguistics*, vol. 13, pp. 78–91, 1947.

[2] E. Selkirk, "The Syllable," in *The Structure of Phonological Representations. Part I*, H. van der Hulst and N. Smith, Eds. Dordrecht: Foris, 1980, pp. 337–383.

[3] M. Jessen, "A survey of German word stress," in *AIMS*, University of Stuttgart, 1995, vol. 2, no. 2, pp. 115–139.

[4] C. Browman and L. Goldstein, "Some notes on syllable structure in articulatory phonology," *Phonetica*, vol. 45, pp. 140–155, 1988.

[5] D. Byrd, "C-Centers revisited," *Phonetica*, vol. 52, pp. 285–306, 1995.

[6] C. Browman and L. Goldstein, "Competing constraints intergestural coordination and self-organization of phonological structures," *Bulletin de la Communication Parlée*, vol. 5, pp. 25–34, 2000.

[7] S. Marin and M. Pouplier, "Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model," *Motor Control*, vol. 14, no. 3, pp. 380–407, 2010.

[8] M. Pouplier, "The gestural approach to syllable structure: Universal, language- and cluster-specific aspects," in *Speech Planning and Dynamics*, S. Fuchs, M. Weihrich, D. Pape, and P. Perrier, Eds. Peter Lang, 2012, pp. 63–96.

[9] J. Pierrehumbert and R. Nair, "Word Games and Syllable Structure," *Language and Speech*, vol. 38, no. 1, pp. 77–114, 1995.

[10] L. Bucar-Shigemori, U. Reichel, and F. Schiel, "Predictability of the effects of phoneme merging on speech recognition performance by quantifying phoneme relations," in *Proc. ESSV*, ser. Elektronische Sprachverarbeitung 2013. Studientexte zur Sprachkommunikation, P. Wagner, Ed. Bielefeld: TUDpress, Dresden, 2013, pp. 247–253.

[11] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo: Morgan Kaufmann, 1993.

[12] R. Baayen, R. Piepenbrock, and L. Gulikers, "The CELEX Lexical Database. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA." CD-ROM, 1995.

[13] C. Browman and L. Goldstein, "Gestural syllable position effects in American English," in *Producing Speech: Contemporary issues: for Katherine Safford Harris*, F. Bell-Berti and R. Lawrence, Eds. American Institute of Physics: Woodbury NY, 1995, pp. 19–34.

[14] J. Bybee, *Phonology and language use*. Cambridge: Cambridge University Press, 2001.

[15] D. Silverman, "Usage-based phonology," in *Continuum Companion to Phonology*. New York City: Continuum, 2011, pp. 369–394.

[16] K. Johnson, "Speech Perception without speaker normalization: An exemplar model," in *Talker Variability in Speech Processing*, K. Johnson and J. Mullenix, Eds., 1997, pp. 145–165.

[17] A. Schweitzer and B. Möbius, "Exemplar-based production of prosody: Evidence from segment and syllable durations," in *Proc. Speech Prosody*, Nara, Japan, 2004, pp. 459–462.

[18] M. Walsh, H. Schütze, B. Möbius, and A. Schweitzer, "An exemplar-theoretic account of syllable frequency effects," in *Proc. Int. Congress of Phonetic Sciences*, Saarbrücken, Germany, 2007, pp. 481–484.

[19] "Clarin-d web page," http://eu.clarin-d.de/index.php/en/.