INTERSEGMENTAL COHESION AND SYLLABLE DIVISION IN POLISH

Pier Marco Bertinetto°, Sylwia Scheuer, Katarzyna Dziubalska-Kołaczyk*, Maddalena Agonigi°*

Scuola Normale Superiore, Pisa°, Adam Mickiewicz University, Poznań*
<Bertinetto@sns.it> <scsylwia@ifa.amu.edu.pl> <dkasia@ifa.amu.edu.pl> <agonigi@sns.it>

ABSTRACT

An experiment with Polish participants was run to shed light on 'intersegmental cohesion hierarchy', with special regard to CC clusters. This hierarchy regulates the strength of the segments' mutual attraction, obeying both universal and language-specific tendencies. The results show that Polish speakers, as contrasted to Italian ones, exhibit a finer cohesion scale due to the richer phonotactics to which they are attuned. In the authors' approach, syllabic structure is assumed to emerge as an epiphenomenon from this hierarchy.

Keywords: syllable, phonotactics.

1. INTRODUCTION

Word games are an established tradition in (psycho-)linguistics. They imply the application of specifically devised alterations to (pseudo-)words, chosen in order to shed light on a particular phonological behavior. The rationale is to gauge the degrees of difficulty that participants meet in applying the same game to different stimuli. Syllable structure has been intensively studied in a number of languages: English, Italian, Spanish, German, Finnish, Korean (cf. [1] for references, [5] for a cross-linguistic comparison).

This paper presents an experiment concerning intersegmental cohesion in Polish, i.e. the degree of attraction between adjacent consonants. Polish was chosen for its highly complex syllable structure. There exist different cohesion coefficients for the various intervocalic CC sequences. This is reminiscent of the traditional notion of 'sonority scale', although the two notions should not be confounded. The 'sonority scale' should be seen as a perceptual-articulatory effect, brought about by manner of articulation (MoA), place of articulation (PoA) and voicing distance (Lx). Intersegmental cohesion crucially depends on the relationships between particular distances [6, 8, 9] and thus results from the complex interplay of adjacent segments, as allowed by universal tendencies as well as by the language-specific phonotactics. Most importantly, intersegmental cohesion determines syllable structure, rather than being determined by the latter.

2. METHODS AND MATERIALS

The materials were (phonotactically legitimate) nonsense disyllables. The three 'games' consisted in syllables inversion or repetition. In the instructions, the word 'syllable' was avoided, in order not to bias the the participants' responses. Each task comprised a training phase, where the participants heard a stimulus immediately followed by the intended alteration. In the test phase, the participants themselves had to apply the alteration, by pronouncing the altered stimulus. The responses were recorded for further analysis.

The disyllabic training items presented single intervocalic consonants. The test list contained 'recall' items (coinciding with the training ones), plus test items with the target CC clusters. These could be modified in more than one way. Consider the item *gopli* in the syllable inversion task. The answer can be either •••***pl***i*go* or •|•*li**g**o**p***, depending on whether the cluster is PRESERVED or SPLIT, as iconically shown by the symbols. Similarly, with syllable repetition one may have: ••*go**g**o**pl**i*, •|•*go**p**go**pl**i*; ••*go**pl**i**pl**i*, •|•*go**pl**i**l**i*. The distribution of the response types should be seen as a measure of intersegmental cohesion, rather than as a syllabification procedure in the strict sense: •• indicates that C_1 is more strongly attracted by C_2 than by the preceding V (thus acting as a word- and syllable-initial cluster); •|• shows the reverse.

Exploiting Polish phonotactics, we used the following five CC cluster classes, all embedded in a CVCCV frame: (1) **OL** 'obstr.+liquid' [subclasses **PL** 'plosive+liquid' and **FL** 'fricat. +liquid']; (2) **LO** 'liquid+obstr.' [subclasses **rO** '/r/+obstr.', **lO** '/l/+obstr.']; (3) **NC** 'nasal+C [subclasses **NC+** (legal Polish initials), **NC-** (illegal Polish initials)]; (4) **FO** 'fricat.+obstr.' [subclasses **FO=** (usually morpheme-internals), **FO≠** (usually across a morpheme boundary), **sC** 'dent. fricat.+C]; (5) **GO** 'glide+obstr.' [subclasses **wO** '/w/+obstr.', (b) **jO** '/j/+obstr.'].

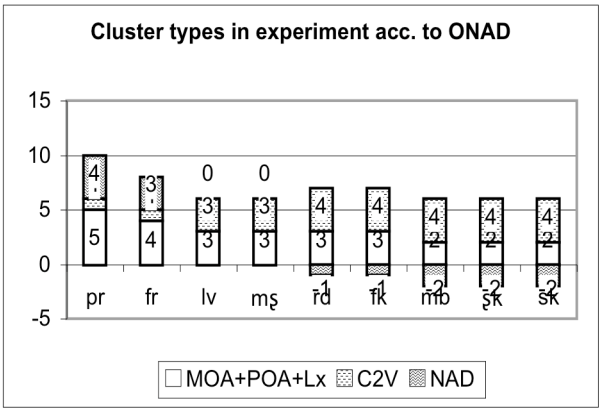
We used 12 stimuli in each subclass, the same for all tasks. All sequences are legitimate medial clusters in Polish. With the exception of NC- and jO (contrasting with the phonotactically legal NC+ and wO), all the clusters can also occur word-initially, which in principle makes •• a viable option. But although the possibility of occurring word-initially is widely accepted as evidence for tautosyllabicity, it is reasonable to assume that not all phonotactically legal word-initial clusters display an equal cohesion degree. This may be influenced not only by universal preferences, but by language-specific factors such as frequency. Relying on [7], the Polish word-initial CC clusters may be arranged in a decreasing frequency order: FO > OL > NC > LO, GO. Actually, FO is not strictly comparable with the other classes, due to its enormous diversity and to a variety of possibly distorting factors: presence/absence of morpheme boundaries, special status of /s/, etc. The actual position of FO in the above hierarchy is thus debatable (see also sect. 5).

Beats-and-Binding phonotactics [8,9] predicts universal cluster preferences in all word-positions. These preferences specify the optimal shape of a particular cluster in a given position by referring to the 'Net Auditory Distance' Principle (NAD) [10,11]. This is a sum of distances in manner of articulation (MOA), place of articulation (POA) and voicing (Lx) between the consonants in a cluster, or a consonant and the neighboring vowels. Thus, NAD = MOA+POA+Lx. A given distance (1, 2, 3 etc.) is assumed between manners (e.g., 1 between stop and fricative), places (e.g., 1 between labial and coronal) and the two major glottis states (1 between vd and vless). The greater the NAD, the better perceptual contrast occurs between the sounds in a sequence. A cluster can only be sustained in a language, if the universal CV preference is counteracted. For instance, the preference relating to initial C_1C_2 clusters takes the form of the following well-formedness condition:

C_1C_2V: NAD $(C_1,C_2) \geq$ NAD (C_2,V)

In prose: In word-initial CC clusters, the NAD between C_1 and C_2 should be greater than or equal to the NAD between a C and a V neighbouring on it, so that the 'bond' tying the two initial Cs counteracts the preferred CV sequence, preserving the cluster. Figure 1 illustrates the NAD values for randomly selected clusters representing the five classes (including some of the eleven subclasses) used in this study. Only two of these clusters qualify as optimal initials and are expected to be preserved on universal (plus language-specific) grounds by the experiment participants.

Figure 1. Selected Polish clusters and the NAD Principle.



PREVIOUS RESULTS

A word games experiment on Italian materials [4], yielded sharp indications. Table 1 presents a results' selection. The CC clusters under scrutiny were: **OL**= 'obstr.+liquid', **LO**= 'liquid+obstr.', **GC**= 'glide+C', **sC** (self-explaining). •• and •|• stand for 'cluster PRESERVED / SPLIT'. Errors were statistically irrelevant; they may be computed for each task by inspecting the horizontal sums. The statistical analyses exploited the Wilcoxon test: * and ** stand for 'significant at the .05 / .01 level', respectively, and accompany the prevailing response type. E.g., in task 2, •• is significantly preferred for OL and •|• for LO and GC.

Table 1: Italian experiment (% data). See text for details.

	1st syll. repetit.		2nd syll. repetit.	
	••	•|•	••	•|•
OL	99,1 **	0	93,3 **	1,2
sC	70,0 *	26,6	41,2	58,3
LO	59,6	39,5	15,0	83,7 **
GC	50,8	45,0	0,4	92,0 **

Over and above the different results yielded by the two tasks (quite unsurprisingly, considering their metalinguistic nature), Italian CC clusters imposed a fairly consistent syllabification strategy, inspired by universal preferences. OL (possible word-initials and tautosyllabic in Italian) sharply contrasted with LO and GC (impossible initials and heterosyllabic), while sC occupied an intermediate position. Indeed, the latter clusters are syllabically undecidable in Italian [2, 3].

PRESENT RESULTS

The results of the Polish experiment appear in table 2: * and ** are as above, + stands for 'close to significance' (.05 ~ .07). The 10 training stimuli were reused as control ones, interspersed as 'recalls' within the task lists. This provided a proficiency measure: of our 12 subjects, one (n.3) was discarded from the statistics of tasks 1 and 2, because of insufficient recall performance.

Consider tasks 1 and 2. While repetition of either the first or the second syllable produced overall differential preferences with the Italian clusters, this happened to a much lesser extent with the Polish ones. Task 1 showed a tendential PRESERVE inclination; task 2, by contrast, yielded a strong SPLIT advantage. Polish speakers tended to repeat the first two segments in task 1 and (to an even larger extent) the last two segments in task 2; they exhibited a parsimonious behavior, consisting in repeating the shortest sequence. Polish clusters are thus relatively plastic (as opposed to the inherently rigid Italian ones). Depending on the task, most of them may be either preserved or split, so that they can by and large alternate between tauto- vs. heterosyllabicity. This is consistent with the fact that Polish phonotactics is, in comparison to Italian, very flexible: the speakers are attuned to a variegated range of consonant sequences and have developed the appropriate articulatory strategies. Hence, the 'bonds' tying adjacent segments are altogether rather flexible, in comparison to the rigid ones operating in the much simpler Italian phonotactics.

Polish clusters are, however, not all alike, as revealed by task 3. The syllables inversion forced the speakers to make a sharp choice (PRESERVE vs. SPLIT). OL and (to a lesser extent) FL exhibited a tautosyllabic inclination, while NC-, lO, rO, jO and wO showed the opposite, heterosyllabic tendency. The remaining clusters (sC, FO≠, FO= and NC+) did not exhibit a statistically interpretable trend.

As for the participants' behavior, there was some variability. Apart from task 2, where •|• significantly prevailed with all participants as a general across-classes trend, tasks 1 and 3 showed a fair amount of individual differences, statistically ranging from no overall preference to •• or •|• preference. Thus, our results stem from averaging over a sample population and by no means reflect the intuitions of every Polish speaker, as should be expected with metalinguistic tasks such as these.

Table 2: Polish experiment (% data). See text for details.

	Task 1 1st syll. repetit.		Task 2 2nd syll. repetit.		Task 3 Syll.s inversion	
	••	•|•	••	•|•	••	•|•
PL	80,30**	15,15	39,39	58,33	63,19 *	13,19
FL	77,27 *	21,97	22,73	76,52 *	63,88 +	20,14
sC	78,03 *	17,42	8,33	91,67**	50,00	31,25
FO=	75,76 +	20,45	14,39	83,33**	49,30	30,56
FO≠	74,24 *	25,00	6,06	93,18**	31,95	51,39
NC+	75,76 *	21,97	10,61	87,12**	29,17	50,00
NC-	68,18 +	25,00	0,75	91,67**	8,34	70,84**
rO	68,18 +	26,72	0	92,42**	18,75	60,42 *
lO	69,70	28,03	0,75	96,97**	15,28	70,14**
jO	68,70 +	25,19	1,51	96,21**	5,56	77,08**
wO	69,70	28,79	1,53	97,71**	17,36	64,58 *

3. DISCUSSION

Comparing the Italian and Polish results may not be straightforward, since more diverse cluster classes were considered in the latter case, due to richer phonotactics. Moreover, task 3 was missing in the Italian experiment. The participants' behavior was, however, sufficiently clear as to make this additional type of evidence unnecessary. Since, in addition, Italians and Poles substantially converge in the treatment of the shared clusters, the present experiment may be interpreted as a contribution to a finer intersegmental cohesion hierarchy, filling the gap as for the clusters not appearing in Italian. Needless to say, over and above the universal tendencies, one should consider language-specific ones: in particular, Italians and Poles seem to diverge as for the treatment of sC clusters. With this in mind, we propose the following generalizations.

Let us first look at tasks 2. OL are obvious onset clusters for Italians, sC less obviously so, while LO and GC are definitely split. Hence, "OL > sC > GC, LO" emerges as the intersegmental cohesion scale for Italian speakers. For Poles, instead, most sequences are split; only PL appears to be a sufficiently good word- and syllable-initial cluster, specially considering the results of task 1. By combining both tasks, "PL > FL > most CC > lO, wO" stands out, as a first approximation, as the Polish cohesion scale. Obviously, for both Italian and Polish one should mention CV as the preferred word- and syllable-initial sequence; however, our experiment only addressed the cohesiveness of intervocalic CC clusters. The strong inclination of the Polish participants towards repeating the first

or (respectively) the last CV sequence in tasks 1-2 merely indicates that most Polish CC clusters are flexible enough to be treated as either indivisible or divisible units, depending on the task.

Task 3 induced further distinctions in the Polish speakers' behavior. OL were kept together; NC-, LO and GO were split, while sC, FO and NC+ remained undecided. Hence, a finer scale emerges: "PL > FL > sC, FO, NC+ > NC-, LO, GO". With respect to figure 1, showing the universally preferred scale of selected clusters, the only modification stemming from this experiment concerns the relative ordering of /mb/ and /fk/ sequences. The only true word- and syllable-onset clusters (/pr/, /fr/) rank in the first two places. Then come legal sequences of low type frequency (/lv/, /mʂ/), still qualifying as possible initials. Finally, we find some universally disfavored initial (and preferred medial) clusters according to NAD, among which our participants: (i) split or treated as undecidable the 'fricat.+stop' clusters (/fk/, /ʂk/, /sk/), which word-initially mostly stem from morphonotactic operations in Polish, and (ii) definitely split /rd, mb/ (the former a very rare initial in Polish, the latter an illegal initial). The type GO, absent in the figure, was always split in the experiment, being either illegal (jO) or a very rare initial (wO).

Checked after [12], the above order corresponds to the clusters frequency. PL constitute over 14% of all #CC- clusters, FL over 4%, still much above 'liquid+fricat.' (0.08%), 'nas.+fricat.' (0.03%) and 'liquid+plos.' (0.01%). The high frequency of 'fricat.+plos.' (over 12%), matching neither the NAD nor our participants' hierarchy, can be explained in Polish by its morphonotactic character, ignored in all available cluster counts.

Summing up, the Polish participants' treatment of CC clusters is governed by: (a) language-specific phonotactics (illegal initial clusters are split), (b) language-specific morphonotactics (morpheme boundaries decidedly win over frequency), (c) universal phonotactic preferences (supporting the major divide into preferred and disfavored initials).

4. CONCLUSION

The cohesion scales for Italian and Polish bear strong resemblances, supporting the universal phonotactic scale on the one hand and the epiphenomenal conception of the 'emergent' syllable on the other one [8,10]. The latter should best be viewed as the result of deeper phonotactic forces, rather than as a structuring phonological primitive. As expected, Polish, with its larger variety of permissible consonant clusters, allowed a finer inspection of the intersegmental cohesion hierarchy. While OL clusters behaved as preferred word- and syllable-onsets in both languages, the remaining clusters exhibited a varying degree of propensity to be treated as onsets. Indeed, with few exceptions (jO, NC- and, to a large extent, FO≠), most of the experimental clusters may also occur word-initially in Polish.

5. REFERENCES

[1] Bertinetto, P.M. 1999a**.** Psycholinguistic evidence for syllable geometry: Italian and beyond. In: Rennison, J. & Kühnhammer, K. (eds), *Phonologica 1996. Syllables!?*. The Hague: Holland Academic Graphics, 1-28.

[2] Bertinetto, P.M. 1999b**.** La sillabazione dei nessi /sC/ in italiano: un'eccezione alla tendenza "universale"?. In: Benincà, P., Mioni, A., Vanelli, L. (eds), *Fonologia e morfologia dell'italiano e dei dialetti d'Italia*. Atti XXXI Congr. SLI. Roma: Bulzoni, 71-96.

[3] Bertinetto, P.M. 2004. On the undecidable syllabification of /sC/ clusters in Italian: Converging experimental evidence. *Italian Journal of Linguistics*16, 349-372.

[4] Bertinetto, P.M., Caboara, M., Gaeta, L., Agonigi, M. 1994. Syllabic division and intersegmental cohesion in Italian. In: Dressler, W.U., Prinzhorn, M., Rennison, J.R. (eds), *Phonologica. Proc. of the 7th International Phonology Meeting*. Torino: Rosenberg & Sellier, 19-33

[5] Derwing, B., Cho, S.W., Samuel, W.H.. 1991. A cross-linguistic experimental investigation of syllable structure: some preliminary results. In: *Proc. 12th ICPhS, Aix-en-Provence*, Vol.4, 110-113.

[6] Dressler, W.U., Dziubalska-Kołaczyk, K. 2006. Proposing morphonotactics. *Wiener Linguistische Gazette* 73, 1-19.

[7] Dunaj, B. 1985. *Grupy Spółgłoskowe Współczesnej Polszczyzny Mówionej* (w Języku Mieszkańców Krakowa). Kraków: Uniwersytet Jagielloński.

[8] Dziubalska-Kołaczyk, K. 2002. *Beats-and-Binding Phonology*. Frankfurt am Main: Peter Lang.

[9] Dziubalska-Kołaczyk, K. 2003. On phonotactic difficulty. *Proc. 15th ICPhS*, *Barcelona*, 2729-2732.

[10] Dziubalska- Kołaczyk, K. Forthc. Universal phonotactics. Studies in Polish Linguistics.

[11] Dziubalska- Kołaczyk, K. & Krynicki G. Forthc. Universal phonotactic calculator: B&B phonotactics and morphonotactics at work. Poznañ Linguistic Meeting 2007.

[12] Śledziński, D.. 2005. *Indeks zbitek spolgloskowych jezyka polskiego z przykladami*. (Index of Polish consonant clusters with examples). Unpublished MS. Poznan: Dept of Ling., A. Mickiewicz University.