

Introduction

Chiara Celata & Basilio Calderone

Phonotactics refers to the principles according to which languages allow sound combinations and segment sequencing to form larger units such as syllables and words. In the study of phonotactics, we are faced with a series of apparent contradictions and empirical problems that require critical comparisons of alternative explanatory models and, most often, an investigation of the ‘interfaces’ between phonotactics and other levels of linguistic organization, particularly phonetics and morphology. One problematic aspect is due to the fact that phonotactics is part of the phonological grammar of a language, and at the same time it is regulated by a number of non-categorical, probabilistic constraints and preferences. It is thus not surprising that the awareness among linguists regarding the role of probability, so crucial in accounting for changes and variations across languages and historical stages (Bod *et al.* 2003), has developed early in connection with observations on the variability in the ‘phonotactic grammar’ of speakers (e.g. Scholes 1966) and on the changing degrees of ‘acceptability’ of word-sized strings (later called ‘wordlikeness’ – a term that explicitly presupposes a probabilistic view of the phonology). A second challenging issue related to phonotactics has to do with the universal versus language-specific nature of phonotactic rules and preferences. Asking what is common to all linguistic systems and what, by contrast, is implemented in individual phonologies under specific conditions has promoted the adoption of a variety of empirical methodologies ranging from the survey of big samples of languages to the psycholinguistic study of how phonotactic structures are processed and acquired, and from probability computations to the investigation of how consonantal and vocalic sequences are produced and perceived.

This paper collection originated as the Second International Workshop on Phonotactics and Phonotactic Modeling (PPM 2013), which was held in Pisa, Italy in November of 2013 and jointly organized by the Scuola Normale Superiore di Pisa and the Université de Paris Ouest.

The articles collected here are a selection of the papers that were presented at the workshop; we warmly acknowledge the authors for their enthusiasm in engaging in this project, their intellectual work

on the initial proposals, and their patience in complying with the many requests from reviewers and guest editors.

This volume can be considered a continuation of the 2014 thematic issue of *Language Sciences* (Elsevier) “Theoretical and empirical approaches to phonotactics and morphonotactics” (Calderone, Celata & Laks 2014), which originated during the First International Workshop on Phonotactics and Phonotactic Modeling (Cortona, Italy) and the thematic session of the 43th Poznań Linguistic Meeting 2012 “Theory and evidence in the study of phonotactics”. It therefore testifies to the living interest in the cross-linguistic, empirically grounded study of phonotactics and in its interfaces. As this preface is being written, the call for papers for PPM 2015 - Third International Workshop on Phonotactics and Phonotactic Modeling (University of Vienna, 26th-27th November, 2015) has just been circulated.

The presentations that were given at PPM 2013 were centred around three thematic sessions: “Phonotactics and frequency, phonotactic modelling”, “Phonotactics and language acquisition”, and “Phonotactics and speech production”. Although the three issues of frequency, acquisition and production are indeed related to each other and most papers deal with more than one issue at a time, the organization of the present volume can be seen as broadly following that thematic partition.

The first three papers deal with cluster complexity and frequency from a typological perspective. The paper by Stefan A. Frisch (University of South Florida) examines the distribution of sonority sequencing (i.e. the preference for clusters having larger steps in sonority difference toward the peak of the syllable) and sonority modulation (i.e. the preference for clusters having larger steps in sonority difference, regardless of sequencing) in a sample of forty-seven languages, and in the phonology of Spanish (according to the frequency count of Spanish medial clusters in the lexicon). Both preferences have been said to be articulatorily or perceptually motivated (e.g. Lindblom 1983, Ohala & Kawasaki-Fukumori 1997), and as such, are expected to be typologically found across languages as well as quantitatively represented in the frequency with which specific clusters occur within a language. The hypothesis is confirmed by the typological survey and the frequency analysis of Spanish clusters, and the paper thus demonstrates the probabilistic nature of functionally based phonological restrictions.

The paper by Katarzyna Dziubalska-Kołaczyk (University of Poznań) discusses implications of the NAD model of phonotactic combinations (Dziubalska-Kołaczyk 2009, 2014) for the theory of mark-

edness in consonant cluster acquisition. It shows that markedness, frequency and age of acquisition are not always straightforwardly correlated; language-specific systemic restrictions must be included to explain both the developmental paths attested in children's acquisition and the statistical biases of cluster occurrence in current use. Independent predictions based on markedness seem therefore to be more consistent with the phonology of languages than with corpus data. These observations open the question of whether it is either the type or the token frequency of phonotactic structures that is relevant for the speakers in the construction of phonotactic grammars and during lexical and morphological acquisition, speech segmentation, phonological categorization etc. (Richtsmeier *et al.* 2010, Calderone *et al.* 2014).

Again from a corpus-based perspective, the paper by Wolfgang U. Dressler, Miroslava Hlinicanová, Matej Durco, Karlheinz Mörth and Katharina Korecky-Kröll (University of Vienna & University of Banská Bystrica) explores type and token frequencies of Slovak and German word-initial, word-medial and word-final consonant clusters (both monomorphemic and across morphemes). It shows that the two languages differ to the extent that German has a proliferation of word-final clusters while Slovak is much richer in word-initial clusters. This difference is related to historical phonology as well as to different morphological processes characterizing the two languages, thus reinforcing the idea of phonotactics as cueing morphological information, in addition to the phonological one. The paper also shows that type and token frequency do not provide entirely overlapping figures, with token frequencies radicalizing the asymmetries found in type frequency. The specificity of adult child-directed speech and its relevance for the study of phonotactic and morphological acquisition is also discussed. Both Dziubalska's and Dressler *et al.*'s papers invite critical reflection upon the methodological choices related to the selection, collection and analysis of different linguistic materials (corpus data, dictionaries, adult speech, child speech etc.), and the impact that such choices may have upon the generalizations that linguists may advance concerning the phonotactics of different languages (or different varieties of the same language).

The two subsequent papers are particularly concerned with language acquisition. The first one, by Laila Kjærbaek, Ditte Boeg Thomsen, Claus Lambertsen and Hans Basbøll (University of Odense) analyzes the early lexical acquisition of Danish by two twin sisters from nine to twenty-nine months old. The focus is on how children increase the length and complexity of their first words when

acquiring the Danish language, which provides very few cues regarding syllable structure and which bears a monotonous sonority envelope – a consequence of its historical development. Making reference to Basbøll's (2005) Sonority Syllable Model, the authors show that in the course of acquisition the two children increase not only the words' length – as calculated by the number of segments and syllables – but also, and perhaps more significantly, the words' phonotactic complexity – as calculated in terms of number of sonority types (i.e., the five major classes of segments as predicted by the Sonority Syllable Model, from vocoids to obstruents with spread glottis) and sonority rises in each produced word.

The paper by Chiara Celata, Katharina Korecky-Kröll, Irene Ricci and Wolfgang U. Dressler (Scuola Normale Superiore di Pisa and University of Vienna) reports on two behavioural experiments testing the reactions of Austrian adults and adolescents to word-final homophonous clusters in intra-morphemic versus cross-morphemic position. The analysis starts from the prediction (contained in Dressler & Dziubalska-Kołaczyk 2006, Dressler *et al.* 2010) that morphonotactics, i.e. the speakers' knowledge of the phonotactics of morpheme concatenation, can affect the processing and acquisition of lexicon, morphology and phonology of a language. The data in fact show that the speakers are either facilitated (in production), or inhibited (in perceptual monitoring) by phonological sequences containing a morpheme boundary, as compared to homophonous monomorphemic sequences. Younger speakers tend to be more sensitive than adults to the morphonotactics of consonant clusters. Thus the results indicate that in processing the phonotactic structures of their native language, the speakers use information from where they are located not only in the word (initial, medial, final) but also in the morphemic structure of the word (within one morpheme or across a morpheme boundary).

The two final papers of the collection deal with the interface between phonotactics and phonetics. The paper by Sylvia Moosmüller (Acoustics Research Institute, Austrian Academy of Sciences) investigates resyllabification in two varieties of Austrian German, namely, Standard Austrian German and the Viennese dialect. The paper highlights differences between the Austrian varieties of German, showing resyllabification before unstressed vowels, and Standard German German, in which resyllabification of word-final consonants is blocked by glottalization. Resyllabification is a syllable-optimizing process generating CV structures and is cued, at the phonetic level, by the temporal reduction of the consonant being produced as syllabic onset. The paper thus demonstrates how changes in the phonotactic organi-

zation of sound sequences impact fine-grained aspects of segment production such as their absolute and relative temporal properties and, finally, are interwoven with the rhythmical output of languages.

The final paper, co-authored by Nathalie Vallée, Thi Thuy Hien Tran, Paolo Mairano, Silvain Gerber & Solange Rossato (CNRS & Université Grenoble Alpes, and Turin University), seeks to provide a phonetic basis for the general phonotactic principle of maximization of syllable onsets and minimization of codas by investigating the jaw oscillatory movements in the repetition of different types of syllables in French and Vietnamese. The results indeed show that the opening phase is generally shorter, faster, and of smaller amplitude than the closing phase (with some minor exceptions). The authors also discuss the findings in a cross-linguistic perspective by pointing out that for other languages, previous studies have found that jaw opening and jaw closing gestures do differ in duration, velocity and amplitude, but in the opposite direction (jaw opening lasted longer and was of greater amplitude than jaw closing). Similar observations emphasize the importance of adopting a cross-linguistic perspective and comparing languages with different phonetic and phonological characteristics in order to uncover the motivations of universal phonotactic preferences, as well as the gestural characteristics of vowel-consonant and consonant-vowel sequence production.

We hope that this special issue will provide interesting suggestions for further investigation, including interdisciplinary approaches, within the domain of the typological, cognitive and physical aspects of sound organization in languages, thus contributing to our knowledge of how human speech structures are mentally organized, acquired, and physically produced.

Addresses of the Authors

Chiara Celata, Laboratorio di Linguistics “G. Nencioni”, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56100 Pisa, Italy <c.celata@sns.it>
Basilio Calderone, CLLE-ERSS, CNRS & Université de Toulouse Jean Jaurès, 5 allées Antonio Machado, 31058 Toulouse, France <basilio.calderone@univ-tlse2.fr>

Bibliographical References

- Basbøll Hans 2005. *The Phonology of Danish*. Oxford: Oxford University Press.
- Bod Rens, Hay Jennifer & Jannedie Stefanie (eds.) 2003. *Probabilistic linguistics*. Cambridge, MA: The MIT Press.
- Calderone Basilio, Celata Chiara & Laks Bernard (eds.) 2014. *Theoretical and empirical models of phonotactics and morphonotactics*. Special issue of *Language Sciences* 46.
- Calderone Basilio, Celata Chiara, Korecky-Kröll Katharina & Dressler Wolfgang U. 2014. A computational approach to morphonotactics: evidences from German. *Language Sciences* 46. 59-70.
- Dressler Wolfgang U. & Dziubalska-Kołodziej Katarzyna 2006. Proposing morphonotactics. *Italian Journal of Linguistics* 18(2), 102-144.
- Dressler Wolfgang U., Dziubalska-Kołodziej Katarzyna & Pestal Lina 2010. Change and variation in morphonotactics. *Folia Linguistica Historica* 31. 51-68.
- Dziubalska-Kołodziej Katarzyna 2009. NP extension: B&B phonotactics. *Poznań Studies in Contemporary Linguistics* 45 (1). 55-71.
- Dziubalska-Kołodziej Katarzyna 2014. Explaining phonotactics using NAD. *Language Sciences* 46. 6-17.
- Lindblom Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle William & Marchal Alain (eds.), *Speech production and speech modelling*. Dordrecht: Kluwer. 403-439.
- Ohala John J. & Kawasaki-Fukumori Haruko 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Eliasson Stih & Hakon Jahr Ernst (eds.), *Language and its ecology: Essays in memory of Einar Haugen*. Berlin: Mouton de Gruyter. 342-365.
- Richtsmeier Peter, Gerken Louann & Ohala Diane 2010. Contributions of phonetic token variability and word-type frequency to phonological representations. *Journal of Child Language* 38(5). 951-978.
- Scholes Robert J. 1966. *Phonotactic Grammaticality*. The Hague: Mouton & Co.

A Preliminary Investigation of Quantitative Patterns in Sonority Sequencing

Stefan A. Frisch

Frisch (2004) claims that phonetically or psycholinguistically-motivated sound patterns should be found typologically across languages (qualitatively), as well as typologically within a language or within the lexicon of a language (quantitatively). This claim emerged from the analysis of Obligatory Contour Principle for Place of Articulation (OCP-Place) consonant patterns in Arabic and a number of unrelated languages as a pattern of similarity avoidance. In this paper, a preliminary study is presented that examines sonority sequencing for consonant clusters within individual languages, and within the lexicons of individual languages for consistent quantitative patterns. Sonority sequencing constraints can be seen as phonetically-motivated in articulation or perception (e.g. Ohala & Kawaski-Fukumori 1997, MacNeilage & Davis 2000). Therefore, sonority sequencing patterns should be present not just in cross-linguistic typology but also in the quantitative patterns within individual languages.

1. Introduction

Linguistically significant phonotactic patterns need not be absolute, but can be statistical. Greenberg (1950) examined consonant combinations in the Arabic lexicon and proposed that cases where consonant pairs were significantly underrepresented compared to expected probability can be used to make linguistically significant generalizations. In these types of studies, the type frequency (frequency in the dictionary) rather than the token frequency (frequency of usage) is generally considered the relevant measure (Bybee 1995) and this approach is supported by children's acquisition of phonological patterns (Munson 2001). More generally, the statistical patterns within the lexicon of phonological forms have been referred to as probabilistic phonotactics (e.g. Vitevitch *et al.* 1999).

1.1. Probabilistic phonotactics

The human cognitive system is known to track frequencies of experiences (in language that would be approximated by corpus token frequencies). These experiences are categorized and classified into abstractions at multiple levels (Munson *et al.* 2011). In the case of phonology, experience with language sound structures is used to define phonological

categories (Pierrehumbert 2003). It has been hypothesized that this type of abstraction is also useful in language acquisition to parse words from speech (Saffran *et al.* 1998), assign word internal phonological and morphological structure (Treiman *et al.* 2000, Hay *et al.* 2003), and provide language specific information for bilinguals (Betancourt 2013; Messer *et al.* 2010; Sebastián-Gallés & Bosch 2002)

Beyond phonotactics, frequency has been shown to influence language processing in a variety of ways. Perception and production are influenced by phoneme frequency or word neighborhood density (Luce & Pisoni 1998; Vitevitch *et al.* 1999). Experiment participants appear to have a sense of phonological frequencies that is reflected in metalinguistic judgments of well-formedness of novel nonwords (Bailey & Hahn 2000). Abstract but quantitative patterns are known by language speakers (Frisch & Zawaydeh 2001; Zuraw 2010; Hayes *et al.* 2009) and have been modeled theoretically in statistical detail (Anttila 2007; Coetzee & Pater 2008; Frisch *et al.* 2004, Hayes & Wilson 2008). These patterns can be modeled by relatively unsophisticated connectionist networks as long as the network is forced to make some sort of generalization (Alderete *et al.* 2013).

1.2. Functionally-grounded phonological explanation and quantitative patterns

Linguistic research has often sought functional explanation for the common typological patterns in the world's languages. Early functional explanations were often dismissed due to their unreliable nature within and across languages. For example, since languages have both harmony and dissimilation processes, it would seem implausible that there is a functional constraint either preferring more similar segments or less similar segments. Recent theoretical developments have made these apparent contradictions less relevant. Optimality Theory, for example, contains mechanisms for resolving conflicting constraints and proposes that in the absence of other factors a functionally unmarked form will emerge universally (Prince & Smolensky 2004). In addition, taking a quantitative perspective on these constraints makes it possible to observe functional influences that involve either more subtle distinctions or are not absolute in their effects. For example, in the case of dissimilation versus harmony it would appear that the Obligatory Contour Principle for Place of Articulation (OCP-Place) is pattern of similarity avoidance (e.g. Frisch *et al.* 2004; MacEachern 1999) but that identity or identity along a particular dimension may be a crucial factor (McCarthy 1994; Berent & Shimron 1997).

The probabilistic view has also created the opportunity to

observe a functional phonological restriction where there is no categorical pattern. Similarity avoidance has been found statistically in English and many other unrelated languages. Additionally, interactions between forces may provide an explanation for the relative strength of the pattern within a language. In the case of similarity avoidance it may be that repeated similar consonants are especially difficult to process given the templatic morphology of Arabic languages where these constraints were originally observed qualitatively (Berg & Abd-El-Jawad 1996).

Frisch (2004) claims on the basis of the similarity avoidance constraint that patterns that are difficult to process cognitively, in perception, or in production will be less frequent than patterns lacking such difficulty. Traditionally, language universals are observed by examining patterns across languages for commonalities, statistical distributions, or implicational relationships. If we add statistical lexical patterns to our search for universal forces acting on language then we can expect that functional forces acting on the sound patterns of language should be found at any level of phonological analysis. Alternatively, this patterning can be viewed as explaining systematic patterns in the number of apparently accidental gaps in otherwise permissible combinations (Frisch *et al.* 2004). Examining quantitative data provides a strong test for claims of functional patterning in phonology. If there is a claim that a pattern is functionally-motivated, it should be found quantitatively in degrees to the extent to which forms obey or violate the proposed functional restriction. In the present study, typologically known patterns in consonant clusters based on sonority are examined within languages for quantitative patterning.

1.3. Functionality of sonority

Sonority constraints on consonant clusters propose that there are manner of articulation restrictions on consonant clusters. Traditionally, low sonority classes are obstruents (stops and fricatives) and high sonority classes are resonant consonants and vowels. Researchers vary number of sonority sub-classes that are defined. Consonant clusters are generally preferred that maximize sonority differences (and when extended to vowels, sonority classification can potentially explain common patterns in syllable structure, e.g. Basbøll 1994). As with many phonological phenomena, there are two broad approaches to the functionality of sonority constraints, based in production or perception. From the production side, sonority sequencing may be the result of ease of production when chunking speech into units (Lindblom 1983; MacNeilage & Davis 2000; Redford 2000). The cycle of jaw movement

from opening to closing to opening again provides a production-based framework for transitioning between low sonority speech sounds (jaw closing or raised for obstruents) to high sonority speech sounds (jaw opening or lowered for vowels). Violations of this sequencing require rapid and precise jaw movements that reverse the overall direction of movement, which would be functionally challenging. Alternatively, violations of this sequencing would be the basis for reorganization of the suprasegmental sound structure. In either case, we might expect violating phonological structures to be dispreferred.

From the perception side, sonority modulation may lead to easier perception of cues to speech sounds and phonological structure. Alternations in amplitude from low to high and/or alternations in high versus low dominant frequency spectra would make speech segments and overall syllabic structure easier to identify (Ohala & Kawasaki-Fukumori 1997). It has also been found that vowel and consonant information are recoverable from either the vowel center or the closure/transition portions of the speech signal, and so the redundant encoding of speech information in both the high and low frequency portions of the spectrum makes speech more robust in a noisy environment.

Given these previous proposals for the functional basis of sonority constraints, specific quantitative hypotheses can be generated. For sonority sequencing, consonant cluster types with larger steps in sonority difference toward the peak of the syllable will be preferred. For sonority modulation, consonant cluster types with larger steps in sonority difference will be preferred, regardless of sequencing. In both cases, a quantitative functional analysis proposes that the degree to which a form is functionally good or bad correlates with the frequency of use of that form in phonotactic patterns within a language. If within language patterns do not vary by frequency as a factor of the degree of sonority difference, then the functional quantitative hypothesis is not supported.

2. *Methods*

2.1. *Quantitative lexical typological study*

The present study examined consonant cluster type frequency for 47 languages (see appendix for sources). Consonant cluster type frequency in this study is the frequency of occurring CC clusters within a language compared to the set of possible clusters given the types of C1 and C2 that are used in clusters. Data are aggregated within broad manner of articulation classes in order to compare rates of

occurrence of different types of clusters by sonority. The present study examines initial, medial, and/or final clusters within the language, depending on the language's phonology and the available data. Note that the source must specifically list the attested consonant clusters in order to conduct a quantitative analysis. If a source only specifies that types of clusters are found it is unusable for quantitative analysis as the apparently accidental gaps cannot be examined. The present study examined CC clusters as the statistical analysis of two segment constituents is the simplest and most straightforward case to analyze (see Pierrehumbert 1994, Frisch 1996, 2000 for issues related to statistical analysis of larger constituents).

The analysis procedure can be demonstrated with a relatively simple language example, the Papuan language Abun (spoken in the West Papua region of Indonesia, Berry & Berry 1999). Word initially, Abun only allows C-glide clusters. If sonority influences the occurrence of clusters then we would expect a relatively greater occurrence of clusters with a large sonority difference (e.g. stop-glide) compared to clusters with a smaller sonority difference (e.g. nasal-glide). However, the raw count of clusters can be deceiving. Abun has a greater variety of stop consonants than nasals so a straightforward comparison of the number of clusters is a misleading representation of the phonotactic possibilities. Specifically, Abun has 7 stops, 4 prenasalized stops, 3 fricatives, 3 nasals, and 2 glides. Table 1 shows the actual number of C-glide clusters for each of these groups, the possible number of C-glide clusters given the consonant inventory of the language, and the relative type frequency of C-glide clusters in each group computed by dividing the actual number by the possible number (A/P). The Abun pattern follows the predictions of sonority modulation and sonority sequencing if we look at the relative number of cluster types. Prenasalized stops are presented as a separate category with higher sonority than stops, but if they were grouped together with the plain stops it would not change the findings, it would just make the example slightly more trivial.

Table 1. Abun C-glide clusters.

C1	ACTUAL	POSSIBLE	A/P
Stop	6	$7 \times 2 = 14$	0.43
Prenasalized stop	3	$4 \times 2 = 8$	0.38
Fricative	1	$3 \times 2 = 6$	0.17
Nasal	1	$3 \times 2 = 6$	0.17

A qualitative analysis of the sonority patterns in Abun has no explanation for the lack of use of some consonant clusters within each sonority level. Such apparently accidental gaps are common, but when examined quantitatively can be seen to be not entirely accidental (Frisch *et al.* 2004). Specifically, there is quantitative variation in cases where otherwise expected consonant clusters are not used. This variation depends on the goodness of the combination from the perspective of sonority, and would be otherwise unexplained if these gaps were considered truly accidental. While in a typical phonological analysis, these gaps are not explained phonologically, they are systematic quantitatively.

As a more complicated example, consider coda CC clusters in English with A/P shown in table 2. In this table, individual consonant counts are not included and the quantitative findings are given in a distilled form with the ratio of use of the possible cluster inventory within each grouping.

Table 2. English CC coda clusters.

C1	C2	STOP	FRICATIVE	NASAL	LIQUID
STOP		0.06	0.10		
FRICATIVE		0.06	0.02		
NASAL		0.22	0.25		
LIQUID		0.92	0.69	0.67	0.25

In order to determine the degree to which a language obeys sonority constraints, data like that in table 2 can be examined by comparing the ratio of actual vs. possible clusters in adjacent cells (e.g. stop-stop vs. stop-fricative). These are minimal steps along the sonority scale. In the case of sonority sequencing, for example, we would predict that fricative-stop codas are more frequent than stop-stop codas which are more frequent than stop-fricative codas. Sonority modulation predicts stop-stop codas are less frequent than fricative-stop codas or stop-fricative codas. With bigger steps in the sonority scale, the differences should become even larger. However, in order to minimize the potential influence of a single larger or smaller numerical value (which is a priori equally likely to fit the hypothesis or not) the cells are only compared in adjacent pairs. This is a stronger test of sonority as each step along the scale is required to change quantitatively in a systematic way (increasing for better sonority combinations, decreasing for worse sonority combinations).

The present analysis starts with comparisons based on sonority modulation, which predicts that A/P will increase away from the main diagonal of the table for all cluster positions. This is relatively easy to compute. In the English coda example in table 2, there is one case where the frequency decreases farther from the main diagonal (nasal-fricative at 0.25 greater than nasal-stop at 0.22). There are three cases where clusters with better sonority modulation are unattested adjacent to cells with attested clusters with worse sonority modulation (adjacent to stop-fricative, fricative-fricative, and liquid-liquid). There are also several ties (one of stop-stop and stop-fricative at 0.06 and several where there are no attested clusters). For the analysis, ties where both adjacent cells have no attested clusters will not be counted as data. Ties where adjacent cells both contain clusters will be counted as not violating sonority modulation. In other words, only quantitative reversals in sonority for attested clusters are considered violations of the hypothesized quantitative pattern. In the case of the English coda consonants, there are 13 pairs of adjacent cells where sonority modulation is respected and 4 pairs where it is reversed, so $13/17 = 76.5\%$ of the comparisons respect sonority modulation. In this way, the entire distribution of consonant cluster patterns with respect to sonority can be distilled to a single value representing the degree to which sonority constraints are obeyed quantitatively across the cluster types in the language.

This analysis, while suggestive, is preliminary and should be viewed with caution due to a number of caveats. First, the finding is based on a one-step pairwise comparison measure between adjacent cells, and so does not capture potentially phonologically relevant sub-patterns in specific cases. Second, the data points are not independent in two ways. Within each language, a single data cell is involved in multiple comparisons with adjacent cells. Using pairwise comparisons for adjacent steps on the sonority scale minimizes this confound as much as possible. In addition, for most of the languages there is more than one cluster position analyzed (e.g. onsets and medial clusters, which may not be phonologically independent). Finally, it is not entirely clear what a random or unstructured distribution of consonant clusters would be. In other words, what is the baseline against which a 76.5% sonority modulation rate in English coda clusters should be compared? These questions are only partially addressed within the scope of the present paper.

2.1.1. Findings: Cluster type frequency

Analysis just like that for English coda consonant clusters, using a table of A/P values for consonant cluster combinations in broad

manner of articulation categories, was conducted for each language in the sample for each cluster position (onset, medial, coda) where data were available. Examining the full set of data across all positions and languages, modulation is obeyed in 72% of adjacent cell comparisons. This is based on 93 cluster data sets from the 47 languages sampled. In other words, there is an overall tendency toward sonority modulation across onset, medial, and coda CC clusters in the relative use of consonant cluster types across a variety of languages. There are quantitative violations of the use of consonant clusters according to sonority modulation within languages, but the overall pattern is robust and consistent with the previously examined case of similarity avoidance (Frisch *et al.* 2004).

With a tendency toward sonority modulation established, the analysis can move forward by considering the influence of sonority sequencing. Sonority modulation is in conflict with sonority sequencing in roughly half of the possible clusters. A sonority fall in an onset is good by sonority modulation but poor by sonority sequencing. Similarly, a sonority rise in a coda is good by sonority modulation but poor by sonority sequencing. Therefore, dividing the data into cases where the two possible functional explanations for sonority either agree or are in conflict can test which of these functional explanations provides a better model of the data. However, in line with other cases of functional explanation for phonological patterns, the interaction between these constraints appears to be cumulative. The interaction between sonority modulation and sonority sequencing across the cluster data sets is shown in table 3.

Table 3. Sonority Modulation and Sonority Sequencing.

Position	N	Constraints agree		Constraints disagree	
		Predicted	SD	Predicted	SD
Initial	38	0.76	0.15	0.69	0.16
Final	18	0.79	0.11	0.58	0.24
Medial*	37	0.71	0.16	0.71	0.18

Table 3 shows that adherence to both sonority modulation and sonority sequencing simultaneously is the most common pattern.

* There is no constraint conflict in medial clusters but the clusters are still split into agree/disagree groups for comparison purposes as though they are onsets (assuming an onset bias, Clements 1990).

When these two constraints are in conflict, the quantitative patterns are less robust. Considering quantitative patterns provides a more detailed understanding of the phonotactics of consonant clusters. What might ordinarily be seen as violations of a strictly defined sonority constraint based on either sequencing or modulation is instead seen as a systematic quantitative patterning that respects, for the most part, both constraints.

There is an additional potential impact on the expected quantitative pattern in consonant clusters described so far. This impact involves the overall use of clusters. If a language has a large number of consonant clusters then they cannot all be optimal with regard to sonority modulation or sonority sequencing. This is similar to the cross-linguistic patterning of vowel segments. Vowels generally distribute themselves evenly around the vowel space (Lindblom 1990), but when a language has a large number of vowels, the phonetic distance between vowels is smaller. This can be seen as a functional constraint against large vowel inventories and quantitatively this appears to be the case across languages. Returning to the case of consonant clusters, having an extremely large number of consonant clusters is likely to create clusters that violate sonority constraints. We can examine the density of clusters in the language (treated as an independent factor) as a functional pressure against obeying sonority.

In addition, considering consonant cluster density also aids in answering the question of the random or unstructured rate of consonant clusters that satisfy sonority constraints. The rate of adherence to sonority modulation patterns for random consonant combinations differs depending on the density of clusters in the language. This was determined by Monte Carlo simulations of an artificial language with densities of 0.1, 0.3, and 0.5 (10%, 30%, and 50%). Consonant cluster pairs were created by randomly distributing the actually occurring consonant clusters among the possible clusters from a relatively generic inventory of obstruents and sonorants. Not surprisingly, with higher cluster density, more violations of sonority are found.

The interaction between cluster density and sonority modulation across the typological data as well as in the Monte Carlo simulations is shown in figure 1. Each data point in the figure is the percent of pairwise comparisons between adjacent sonority cells that respect sonority modulation for a particular language and cluster position in the sample (as exemplified for English coda clusters in table 2). A linear regression shows a decreasing tendency to obey sonority modulation as cluster density increases. However, the mean across languages is well above that predicted by chance in the Monte Carlo simulations

for the three different densities (in the vicinity of 50% and decreasing with increasing density). In addition to the density-based pattern in sonority modulation, there is an overall tendency to avoid large numbers of clusters in the languages sampled.

2.1.2. *Summary: Cluster type frequency*

More “gaps” are found in permissible sonority combinations when the sonority difference is small compared to when the sonority difference is large, showing a tendency to obey a sonority modulation constraint. This constraint is more robust when the sonority difference is also compatible with sonority sequencing and less robust when sonority sequencing is violated, showing that both sonority modulation and sonority sequencing influence the cross-linguistic and within language frequency of cluster types. Functional sonority constraints interact with a third factor, in that it is harder to obey sonority restrictions if the language uses many clusters. The sonority constraints are more robust in languages with sparse cluster use and more likely to be violated with dense cluster use. Overall, then, there

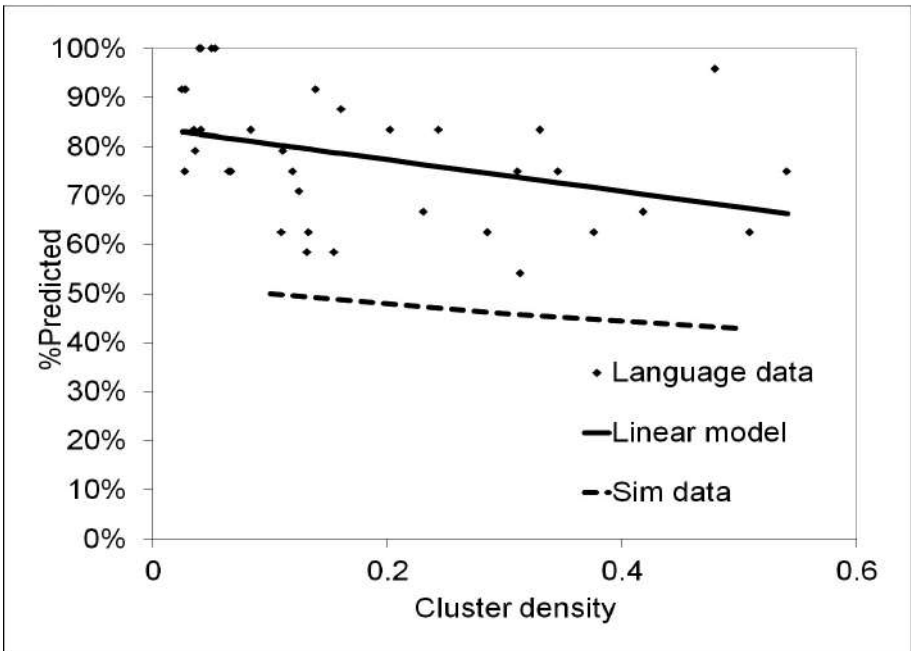


Figure 1. Clusters respecting sonority modulation by cluster density in the language.

are three factors influencing the cross-linguistic distribution of gaps in cluster inventories. Languages avoid large numbers of clusters, avoid clusters that violate sonority modulation, and avoid clusters that violate sonority sequencing. These three constraints interact cumulatively when examined quantitatively.

2.2. Study of cluster lexical frequency

If sonority constraints are functional, grounded in phonetic or psycholinguistic processing principles, then we would predict the functional pressures that lead to an avoidance of sonority violations apply as individual words are used. By definition, functional constraints make the individual words more difficult to perceive or produce correctly. Thus, we would expect to find quantitative patterns related to sonority sequencing within the frequency distribution of forms in the lexicon (Frisch 2004). Quantitatively examining sonority patterns in the lexicon requires consideration of the frequency of occurrence of the consonants involved in the cluster. A cluster involving two infrequent consonants would be unlikely to occur regardless of sonority. A gap in clusters involving two frequent consonants is more likely to be systematic. For the present study, the O/E (observed divided by expected) measure of relative co-occurrence is used to account for baseline segment probability (Frisch *et al.* 2004).

An analysis of cluster frequency in the lexicon has been completed for two data sets, Spanish medial clusters and English coda clusters. The case of Spanish medial clusters will be covered in greater detail. Table 4 shows the frequency count of across lexical items for Spanish medial clusters of different sonority classes in the University of South Florida Spanish Frequency Lexicon (Brea-Spahn 2009; Frisch & Brea-Spahn 2010).

Table 4. Frequency count for Spanish medial clusters in the lexicon.

C1	C2	Stop	Fricative	Nasal	Liquid	Glide
Stop		299	385	70	1214	520
Fricative		1522	137	162	515	1852
Nasal		2575	931	43	16	328
Liquid		578	610	308	17	528
Glide		93	92	65	37	0

Given the frequency of different sonority classes of consonant in clusters, the randomly expected frequency of sonority combinations can be computed by multiplying the marginal probabilities for each type of C1 and C2. The observed count divided by the count expected by chance (O/E) provides a relative measure of co-occurrence given baseline frequency. Table 5 shows the O/E for the Spanish medial data. An O/E less than 1 means that fewer combinations than expected are found. An O/E greater than 1 means that more combinations than expected are found. Like with the analysis of consonant cluster types, we would predict relative frequency to increase away from the main diagonal of the table. Comparing adjacent cells as before, the percent of comparisons that respect sonority modulation can be computed to provide a single number that summarizes the degree to which quantitative patterns obey the hypothesized constraints.

Table 5. O/E for Spanish medial clusters in the lexicon.

C1	C2	Stop	Fricative	Nasal	Liquid	Glide
Stop		0.24	0.77	0.47	1.97	0.69
Fricative		0.94	0.21	0.82	0.63	1.87
Nasal		1.74	1.54	0.24	0.02	0.36
Liquid		0.80	1.76	2.94	0.04	1.01
Glide		0.86	2.04	4.87	0.68	0.00

A similar analysis was conducted for English final CC clusters in the Hoosier Mental Lexicon database of English words (Nusbaum *et al.* 1984). As with the analysis of cluster type frequency, comparisons between two empty cells were not used in the computation.

2.2.1. Findings: Cluster lexical frequency

In the Spanish medial cluster data, 68% of comparisons obey sonority modulation for O/E across the Spanish lexicon. In the English final cluster data, 71% of comparisons obey sonority modulation for O/E across the English lexicon. These findings are similar to the findings for cluster type frequency. This pattern is consistent with a soft or gradient constraint against poor sonority modulation clusters acting on the lexicon of the language, influencing the distribution of consonant clusters in lexical items. Finding the general typological pattern for sonority within the quantitative distribution of lexical items within a language provides strong evidence for the functional nature of the constraints involved.

2.3. Limitations of the present study

The current study is typologically limited to a convenience sample of languages. For the most part, this is a subset of languages from Greenberg (1978). A more extensive study to verify the claims of this paper is needed with a larger and typologically balanced sample of languages. In addition, the present study examined only two cases of patterns within a lexicon. To clearly establish the quantitative nature of sonority constraints in consonant clusters, additional lexical sources will have to be investigated.

The present study is also based on a relatively standard division of manner classes for consonants into a sonority hierarchy. However, this division has been claimed to have no logical basis by Basbøll (1994), who advocates a more generic sonority hierarchy with only three distinctions. Given that the quantitative patterns observed here work reasonably well and are relatively robust, some variation in the definition of the sonority hierarchy will not affect the overall result. This is one advantage of taking a quantitative approach. Minor contradictory details have little impact on the overall pattern.

3. Conclusion

The strongest possible test of functional constraints on language processing would find both quantitative and qualitative patterns in language that reflect functional forces. If a typological pattern is functionally-based, then the cross-linguistic pattern should also be visible within individual languages quantitatively (Frisch 2004). In the present study, articulatory and perceptual motivations for the sonority hierarchy were considered and patterns within languages' cluster inventories and lexicons were examined. Parallel with the previously examined Obligatory Contour Principle for Place of Articulation (OCP-Place) as a functional constraint on similarity avoidance, quantitative patterns were found within languages that reflect typological generalizations in sonority modulation and sonority sequencing. In this preliminary analysis, the hypothesis is supported: Quantitative sonority constraints on consonant clusters are found reflecting potential functional motivations for sonority in both production and perception. The prediction is that functionally-motivated phonotactic patterns in sonority constraints should be found, at least statistically, everywhere we look.

Stefan A. Frisch

Address of the Author

Stefan Frisch, University of South Florida, Dept. of Communication Sciences and Disorders, 4202 E. Fowler Ave, PCD1017, Tampa, FL 33620-8100
<sfrisch@usf.edu>

Bibliographical References

- Alderete John, Tupper Paul & Stefan A. Frisch 2013. Phonological constraint induction in a connectionist network: Learning OCP-Place constraints from data. *Language Sciences* 37. 52-69.
- Anttila Arto. 2007. Variation and optionality. In de Lacy Paul (ed.), *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press. 519-536.
- Bailey Todd & Ulrike Hahn 2001. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 4. 568-591.
- Basbøll Hans 1994. How to derive the sonority syllable from the prototypical peak. *Acta Linguistica Hafniensia* 27. 51-65.
- Berent Iris & Joseph Shimron 1997. The representation of Hebrew words: Evidence from the obligatory contour principle. *Cognition* 64. 39-72.
- Berg Thomas & Hassam Abd-El-Jawad 1996. The unfolding of suprasegmental representations: A cross-linguistic perspective. *Journal of Linguistics* 32. 291-324.
- Betancourt Kyna S. 2013. Nonword processing in bilingual five year olds: Do phonotactics count? Unpublished Ph.D. dissertation, University of South Florida.
- Blevins Juliette 2004. *Evolutionary Phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Blevins Juliette 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32. 117-166.
- Brea-Spahn María R. 2009. Spanish-specific patterns, language proficiency, and nonword repetition performance in English language learners. Unpublished Ph.D. dissertation, University of South Florida.
- Bybee Joan 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10. 425-455.
- Bybee Joan 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14. 261-290.
- Bybee Joan 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- Clements George N. 1990. The role of the sonority cycle in core syllabification. In Kingston John & Mary Beckman (eds.), *Laboratory phonology I: Between the grammar and the physics of speech*. New York: Cambridge University Press. 283-333.
- Coetzee Andries & Joe Pater 2008. Weighted constraints and gradient restric-

- tions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26. 289-337.
- Frisch Stefan A. 1996. Similarity and frequency in phonology. Unpublished Ph.D. dissertation, Northwestern University.
- Frisch Stefan A. 2000. Temporally organized lexical representations as phonological units. In Broe Michael B. & Pierrehumbert Janet B. (eds.), *Acquisition and the lexicon: Papers in laboratory phonology V*. Cambridge: Cambridge University Press. 283-298.
- Frisch Stefan A. 2004. Language processing and OCP effects. In Hayes Bruce, Kirchner Robert & Donca Steriade (eds.), *Phonetically-based phonology*. Cambridge: Cambridge University Press. 346-371.
- Frisch Stefan A. & María R. Brea-Spahn 2010. Metalinguistic judgments of phonotactics by monolinguals and bilinguals. *Laboratory Phonology* 1. 345-360.
- Frisch Stefan A., Pierrehumbert Janet B. & Michael B. Broe 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22. 179-228.
- Frisch Stefan A. & Bushra Zawaydeh 2001. The psychological reality of OCP-Place in Arabic. *Language* 77. 91-106.
- Greenberg Joseph H. 1950. The patterning of root morphemes in Semitic. *Word* 5, 162-181.
- Greenberg Joseph H. 1978. In Greenberg Joseph H., Ferguson Charles A. & Edith A. Moravcsik (eds.), *Universals of human language: Phonology*. Stanford, Stanford University Press. 243-279.
- Hay Jennifer, Pierrehumbert Janet B. & Mary Beckman 2003. Speech perception, wellformedness, and the statistics of the lexicon. In Local John, Ogden Richard & Rosalind Temple (eds.), *Papers in laboratory phonology VI: Phonetic interpretation*. Cambridge: Cambridge University Press. 58-74.
- Hayes Bruce, Kirchner Robert & Donca Steriade 2004. *Phonetically-based phonology*. Cambridge: Cambridge University Press.
- Hayes Bruce & Colin Wilson 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379-440.
- Hayes Bruce, Kie Zuraw, Péter Siptár & Zsuzsa Londe 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85. 822-863.
- Lindblom Björn 1983. Economy of speech gestures. In MacNeilage Peter (ed.), *The production of speech*. New York: Springer-Verlag. 217-246.
- Lindblom Björn 1990. Explaining phonetic variation: A sketch of the H & H theory. In Hardcastle William & Alain Marchal (eds.), *Speech production and speech modelling*. Dordrecht: Kluwer. 403-439.
- Luce Paul. A. & David B. Pisoni 1998. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* 19. 1-36.
- MacEachern Margaret R. 1999. *Laryngeal cooccurrence restrictions*. New York: Garland.
- McCarthy John J. 1994. The phonetics and phonology of Semitic pharyngeals. In Keating Patricia (ed.), *Papers in laboratory phonology III: Phonological structure and phonetic form*. Cambridge: Cambridge University Press. 191-283.

- MacNeilage Peter F. & L. Davis Barabara 2000. On the origin of internal structure of word forms. *Science* 288. 527-531.
- Messer Marielle H., Leseman Paul P.M., Boom Jan & Aziza Mayo. Y. 2010. Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology* 105. 306-323.
- Munson Benjamin 2001. Phonological pattern frequency and speech production in children and adults. *Journal of Speech, Language, and Hearing Research* 44. 778-792.
- Munson Benjamin, Edwards Jan and Mary E. Beckman 2011. Phonological representations in language acquisition: Climbing the ladder of abstraction. In Cohn Abigail C., Fougeron Cécile & Marie K. Huffman (eds.), *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press. 288-309.
- Nusbaum Howard, Pisoni David & Christopher. K. Davis 1984. Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No. 10* (Speech Research Laboratory, Bloomington IN). 357-376.
- Ohala John J. & Haruko Kawasaki-Fukumori 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Eliasson Stih & Jahr Ernst Hakon (eds.), *Language and its ecology: Essays in memory of Einar Haugen*. Berlin: Mouton de Gruyter. 342-365.
- Pierrehumbert Janet 1994. Syllable structure and word structure: A study of triconsonantal clusters in English. In Keating Patricia (ed.), *Papers in laboratory phonology III: Phonological structure and phonetic form*. Cambridge: Cambridge University Press. 168-188.
- Pierrehumbert Janet (2003) Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46. 115-154.
- Prince Alan & Paul Smolensky 2004. *Optimality theory: Constraint interaction in generative grammar*. Oxford: Blackwell Publishers.
- Redford Melissa A. 2000. *An articulatory basis for the syllable*. Unpublished Ph.D. dissertation, University of Texas at Austin.
- Saffran Jenny R., Aslin Richard N. & Elissa L. Newport 1998. Statistical Learning by 8-month-old Infants. *Science* 274. 1926-1928.
- Sebastián-Gallés Nuria, Echeverría Sagrario & Laura Bosch 2005. The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language* 52. 240-255.
- Treiman Rebecca, Kessler Brett, Knewasser Stephanie, Tincoff Ruth & Margo Bowman 2000. English speakers' sensitivity to phonotactic patterns. In Broe, Michael & Pierrehumbert, Janet B. (eds.), *Papers in laboratory phonology V: Language acquisition and the lexicon*. Cambridge: Cambridge University Press. 269-282.
- Vitevitch Michael S., Luce Paul A., Pisoni David B. & Edward T. Auer 1999. Phonotactics, neighborhood activation and lexical access for spoken words, *Brain and Language* 68. 306-311.
- Zuraw Kie 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language and Linguistic Theory* 28. 417-472.

Appendix: Sources

IJAL = *International Journal of American Linguistics*

HAIL = *Handbook of American Indian Languages*

Abramson A. (1962). *The vowels and tones of Standard Thai*. Bloomington, IN: Indiana University Press.

Aschmann H. P. (1946). Totonaco phonemes. *IJAL*, 12, 34-43.

Andrews H. (1949). Phonemes and morphophonemes of Temoayan Otomi. *IJAL*, 15, 213-222.

Berry K. & Berry C. (1999). *A description of Abun: A West Papuan language of Irian Jaya*. (Pacific Linguistics, Series B, 115.) Canberra: Australian National University.

Boas F. & Deloria E. (1976). *Dakota grammar*. New York: AMS press.

Bogoras W. (1922). Chukchee. In Boas F. (ed.), *HAIL* 2. Washington, DC: Government Printing Office. 631-903.

Callister B. (1993). Misiman phonology. In J.M. Clifton (ed.), *Phonologies of Austronesian Languages*, No. 2. SIL, Papua New Guinea.

Camara Jr J. M. & Mattoso J. (1979). *The Portuguese language*. Chicago: University of Chicago Press.

Chitoran I. (1994). Acoustic investigation of Georgian harmonic clusters. *Working papers of the Cornell phonetics laboratory*, 9, 27-65.

Deprez V. (1988). Georgian complex segments. *Proceedings of NELS*, 16, 109-123.

Fast P.W. (1953). Amuesha (Arawak) phonemes. *IJAL*, 19, 191-194.

Ferguson C.A. & Chowdhury M. (1960). The phonemes of Bengali. *Language*, 36, 22-59.

Fractenberg L.J. (1922). In Boas F. (ed.), *HAIL* 2. Washington, DC: Government Printing Office. 297-429.

Garvin P.L. (1950). Wichita I: Phonemics. *IJAL*, 16, 179-184.

Id. (1948). Kutenai I: Phonemics. *IJAL*, 14, 37-42.

Gibson L.F. (1956). Pame (Otomi) phonemics and morphophonemics. *IJAL*, 22, 242-265.

Harden M. (1946). Syllable structure of Terena. *IJAL*, 12, 60-63.

Jaeger J.J. & Van Valin R.D. (1982). Initial consonant clusters in Yatee Zapotec. *IJAL*, 48, 125-138.

Kohler K. (1999). German. In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press. 86-89.

Kordić Snježana (1997). Serbo-Croatian (Languages of the world/materials, 148.) Munich: Lincom Europa.

Koutsoudas A. (1962). *Verb morphology of Modern Greek*. Bloomington, IN: Indiana University Press.

Lisker L. (1963). *Introduction to spoken Telugu*. New York: American Council of Learned Societies.

Marlett S.A. (1988). The syllable structure of Seri. *IJAL*, 245-278.

Matthews G.H. (1955). A phonemic analysis of a Dakota dialect. *IJAL*, 21, 56-59.

McArthur H. & McArthur L. (1956). Aguacatec (Mayan) phonemes in the stress group. *IJAL*, 22, 72-76.

- McIntosh J.B. (1945). Huichol phonemes. *IJAL*, 11, 31-35.
- McKaughan H.P. (1954). Chatino formulas and phonemes. *IJAL*, 20, 23-27.
- Morelli F. (1999). The phonotactics and phonology of obstruent clusters in Optimality Theory. Unpublished Ph.D. dissertation, University of Maryland.
- Morgan T.A. (1988). Positive and negative syllable structure conditions in Spanish. In T.A. Morgan, J.F. Lee & B. VanPatten (eds.), *Language and language use: Studies in Spanish*. New York: University Press of America. 37-54.
- Nakaskul K. (1978). The syllabic and morphological structure of Cambodian words. In P. Jenner (ed.), *Mon-Khmer studies VII*. 183-200.
- Nater H. F. (1984). The Bella Coola language (Canadian ethnology service paper no. 92). Ottawa: National Museums of Canada.
- Needham D. & Davis M. (1946). Cuicateco phonology. *IJAL*, 20, 23-27.
- Newmark L. (1957). *A structural grammar of Albanian*. Bloomington, IN: Indiana University Press.
- Penzl H. (1955). *A grammar of Pashto*. Washington: American Council of Learned Societies.
- Pike K.L. & Pike E.V. (1947). Immediate constituents of Mazateco syllables. *IJAL*, 12, 78-91.
- Powell J.V. & Woodruff Sr F. (1976). *Quileute Dictionary*. Moscow, ID: University of Idaho Press.
- Prick van Wely F.P.H. (1951). *Cassell's English-Dutch, Dutch-English dictionary*. London: Cassell.
- Rabel L. (1961). *Khasi, a language of Assam*. Baton Rouge: Louisiana State University Press.
- Rath D.D. (1993). Mengen phonology essentials. In J. M. Clifton (ed.), *Phonologies of Austronesian Languages, No. 2*. SIL, Papua New Guinea.
- Ruhlen M. (1973). *Rumanian phonology*. Stanford: Stanford University Press.
- Reichard G.A. (1933). Coeur d'Alene. In Boaz, F. (ed.), *HAIL 3*. Washington, DC: Smithsonian Institute. 515-707.
- Robbins F.E. (1961). Quiotepec Chinantec syllable patterning. *IJAL*, 27, 237-250.
- Salzmann Z. (1956). Arapaho I: Phonology. *IJAL*, 22, 49-56.
- Sapir E. (1922). The Takelma language of Southwestern Oregon. In Boas, Franz (ed.), *HAIL 2*, 1-296. Washington, DC: Government Printing Office.
- Siptar P. & Torkenczy M. (2000). *The phonology of Hungarian*. Oxford: Oxford University Press.
- Skinner L.E. (1962). Usila Chinantec syllable structure. *IJAL*, 28, 251-255.
- Spencer R.F. (1946). The phonemes of Keresan. *IJAL*, 12, 229-236.
- Spotts H. (1953). Vowel harmony and consonant sequences in Mazahua (Otomi). *IJAL*, 19, 253-258.
- Steriade D. (1982). Greek prosodies and the nature of syllabification. Unpublished Ph.D. dissertation, MIT.
- Tarpen M.L. (1989). A grammar of the Nisgha language. Unpublished Ph.D. dissertation, University of Victoria.
- Trager F.H. (1971). The phonology of Picuris. *IJAL*, 37, 29-33.

A Preliminary Investigation of Quantitative Patterns in Sonority Sequencing

- Vanvik A. (1972). A phonetic-phonemic analysis of Standard Eastern Norwegian. *Norwegian Journal of Linguistics*, 26, 119-164.
- Waterhouse V. & Morrison M. (1950). Chontal phonemes. *IJAL*, 16, 35-39.
- Whitman W. (1947). Descriptive grammar of Ioway-Oto. *IJAL*, 13, 233-248.
- Wolff H. (1952). Osage I: Phonemes and historical phonology. *IJAL*, 18, 63-68.
- Wonderly W.L. (1951). Zoque II: Phonemes and morphophonemes. *IJAL*, 17, 105-123.
- Wright R. (1996). Consonant clusters and cue preservation in Tsou. Unpublished Ph.D. dissertation, UCLA.
- Zhukova A.N. (1967). Russian-Koryak dictionary. Moscow: Izdat, Sovetskaja encilopedija.

Are frequent, early and easy clusters also unmarked?

Katarzyna Dziubalska-Kołaczyk

Phonological theories generally assume that unmarked clusters would demonstrate similar behaviour in language use, language acquisition and language production. In particular, they are expected to be more frequent, acquired earlier and easier to produce than the marked ones. It seems, however, that the frequent, early and easy clusters are nonidentical sets. Additionally, morphonotactic clusters, which are by default more prone to be marked, behave differently than the marked phonotactic clusters, in particular, they are acquired earlier. In the paper, I will draw from the recent research by myself and colleagues¹ in order to discuss the difficulty in avoiding circularity in characterizing markedness as well as the omnipresent methodological bias connected with the choice of data, both experimental and collected, used to support claims concerning phonotactics.

1. Introduction

Five most frequent initial CC clusters in Polish are /pʃ, pr, st, mj, sp/ (cf. Zygorowicz *et al.* in press). Some studies (Dziubalska-Kołaczyk 1999, Zygorowicz 2010) on first language acquisition of Polish mention /sp, fʃ, vw, st, sw/ among those acquired early, and the morphonotactic² among them as acquired prior to the lexical, i.e. intramorphemic, ones (e.g. lexical /fʃ/ would be reduced much more often than the morphonotactic one). Assessment of all those clusters from the perspective of speech production is a complex task, involving detailed phonetic description focused on coarticulatory effects and gestural coordination.³ Importantly, production-based criteria interact with the acoustic/auditory ones. Additionally, such descriptions are always theory-grounded and therefore provide divergent assessments of 'ease'. Already a shift from a production to a perception perspective leads to conflicting interpretations. For example, a homorganic cluster is 'easier' to produce but 'less easy' to perceive. Despite the above problems of analysis, what can be noticed is that /st/ and /sp/ clusters figure in both of the above sets of the frequent and of the early acquired clusters. It is well known that s+stop clusters are notoriously unaccountable for across models of phonotactic markedness. Sonority-based models generally fail to explain their occurrence. Phonetic, phonological and typological accounts point

to the uniqueness of those clusters (cf. an overview by Olender 2013). Still, phonologically speaking, they are marked, and require a special treatment across models. As already mentioned, they are also frequent and early, possibly also easy.

The above example introduces the topic of this paper: the relationship between independently defined markedness and frequency, order of acquisition and ease of production. The major advantage of the concept of markedness has been to allow for predictions of the occurrence of structures in given types of linguistic behaviour. The less marked (=more natural) a structure, the more likely it would be to occur frequently, to be acquired early and to be easy to produce. Of course, such inference may only work on condition that markedness be defined independently of the criteria of frequency, age of acquisition or ease of articulation. Otherwise, it becomes circular.

In this paper we deal with markedness in phonotactics. Clusters are deemed more or less natural depending on a measure called NAD (Net Auditory Distance). One expects the clusters preferred by NAD to be more frequent, to be acquired before the dispreferred ones as well as to be easier to produce. First, the phonological theory underlying the model of phonotactics will be briefly introduced (section 2) and NAD as a measure of markedness will be presented (section 3). Next, the data illustrating frequency, acquisition and ease will be discussed and assessed from the perspective of NAD (section 4). Section 5 will contain conclusions.

2. *Beats-and-Binding (B&B) phonotactics*

The model of phonotactics operating with NAD (cf. Dziubalska-Kołaczyk 2009, 2014) is part of the syllable-less theory of phonology called Beats-and-Binding (B&B) phonology (Dziubalska-Kołaczyk 2002). The rationale behind this model of phonotactics is to counteract the preference for CV (but cf. Hyman 2008), which overwhelmingly forces clusters to wither either via reduction or epenthesis/prothesis or at least substitution (CCVCV, CCV→CVCV, CCV→VCCV). Clusters of consonants persist thanks to auditory contrast between members of a cluster and its satisfactory distribution across the word. The principle of perceptual contrast is grounded in the semiotic figure-and-ground principle, on the one hand, and in phonetics, on the other, since as observed by Ohala (1990), larger modulations in several acoustic parameters (amplitude, periodicity, spectral shape, F0) have more survival value than lesser ones and therefore will persist in languages.

3. NAD

3.1. Definition of NAD

The Net Auditory Distance (NAD) Principle defines cluster preferability in relation to the position in the word (initial, medial and final). It reads:

A cluster is preferred if it satisfies a pattern of distances specified by the universal phonotactic preference relevant for its position in the word.

For example, for the initial cluster C1C2V:

$NAD(C1, C2) \geq NAD(C2, V)$. This notation means that, in word-initial double clusters, NAD between the two consonants should be greater than or equal to NAD between a vowel and a consonant neighbouring on it.

3.2. Measuring NAD

It is believed (Dziubalska-Kořaczyk 2002, 2009, 2014) that auditory (perceptual) distance can be expressed by combinations of articulatory features, which eventually bring about the auditory effect. An indefinite number of articulatory features as well as detailed acoustic cues would have to be investigated in terms of the degree of their contribution to the overall auditory effect obtained in a cluster. As a starting point, manner of articulation (MOA) and place of articulation (POA) are selected, as well as the S/O (sonorant/obstruent) distinction. Tables 1 and 2 illustrate the way in which number values are being assigned to particular POAs and MOAs. The presence of the S/O distinction is signalled by 1, its absence by 0.

The manners and places assumed in the tables are selected according to their potential relevance: 6 manners (stop, affricate, fricative, sonorant stop, approximant, semivowel), where affricates and semivowels are, tentatively, attributed half a distance due to their ambiguous nature; and 5 places (labial, coronal, dorsal, radical and laryngeal or glottal). Manners refer to the most generally acknowledged version of the so-called sonority scale, while places are taken from Ladefoged (2006: 258). Both lists are extendible and modifiable (e.g., Ladefoged's list consists of 5 nodes which branch into 12 more detailed features), depending on the amount of detail we want to include in the definition of distance. Importantly, vowels need to be differentiated according to their colour, too, which will allow reflecting the intersegmental distances with a higher degree of precision.

Table 1 is meant as a general, universal reference for the values of MOA and POA which constitutes a guideline for the creation of language-specific tables. Table 2 is the adaptation of Table 1 to Polish. Polish has three values for obstruents, two values for liquids, an additional distinction within glides (nasal vs. oral), has a value for labiodental within labials, three values within the feature coronal, and two values within the feature dorsal.

Table 1. Distances in MOA and POA.

4		3		2		1		0	
OBSTRUENT				SONORANT				VOWEL	
STOP		FRICATIVE		SONORANT STOP		APPROXIMANT			
	(AFFRICATE)						semiV		
p b		ɸ β f v		m ɱ		w		labial	1
t d t̪ d̪		θ ð s z ʃ ʒ		n		r l		coronal	2
k g c ɟ		ç ʒ x ɣ		ɲ ŋ		j		dorsal	3
								radical	4
ʔ		h						laryngeal (glottal)	5

Table 2. Distances in MOA and POA: Polish.

OBSTRUENT			SONORANT					VOWEL		
STOP		FRICATIVE	NASAL	LIQUID		GLIDE				
	AFFRICATE			2.0						
				lateral	rhotic					
5.0	4.5	4.0	3.0	2.5	2.0	1.5	1.0	0		
p b			m			ɰ	w	1.0	bilabial	LABIAL
		f v						1.5	labio-dental	
t d ts dz s z			n	l				2.0	(post-) dental	CORONAL
tʃ dz ʃ z				r				2.3	alveolar	
tɕ dz ɕ z			ɲ					2.6	alveolo-palatal	
						ĩ	j	3.0	palatal	DORSAL
k g		x	ŋ			ũ	w	3.5	velar	
								4.0		RADICAL
								5.0		GLOTTAL

Are frequent, early and easy clusters also unmarked?

3.3. An example of a calculation⁴

For an initial Polish cluster /prV/, NAD preferences should be calculated as follows:

$$\text{NAD CC} = |\text{MOA1} - \text{MOA2}| + |\text{POA1} - \text{POA2}| + \text{S/O}$$

$$\text{NAD CV} = |\text{MOA1} - \text{MOA2}| + \text{S/O}$$

$$\text{NAD CC} = |5-2| + |1-2.3| + 1 = 3 + 1.3 + 1 = 5.3$$

$$\text{NAD CV} = |2-0| + 0 = 2 + 0 = 2$$

$$5.3 > 2$$

So, the preference $\text{NAD}(\text{C1}, \text{C2}) \geq \text{NAD}(\text{C2}, \text{V})$ is observed since $5.3 > 2$.

3.4. NAD as a measure of markedness: predictions

NAD allows for the formulation of two types of predictions. On the one hand, it distinguishes between two sets of clusters, preferred (i.e. those which observe the preferences) and dispreferred ones (those which do not), which constitutes a binary qualification into unmarked=preferred=natural vs. marked=dispreferred. On the other hand, it allows to build a hierarchy of preference of clusters ranging from the most preferred to the most dispreferred (least preferred), which constitutes a scalar qualification of relative markedness from the most natural=unmarked=preferred to the heavily marked. This clearly is a strong attribute of this measure, since it reflects the human ability to both categorize and build continua. The matrices below (Figure 1, 2 and 3, prepared by Michał Jankowski) provide an

	p	b	t	d	k	g	ts	dz	tf	dz	tf	dz	f	v	s	z	f	z	ɕ	ʂ	x	n	ɲ	m	ɲ	l	r	w	j	ɰ	ɰ
p			NO						NO						NO												NO				
b																															
t									NO					NO											NO		NO	NO	NO		
d																															
k			NO												NO											NO		NO			
g																															
ts																															
dz			NO																												
tf	NO		NO																												
dz																															
tf																															
dz																															
f																															
v																															
s	NO		NO		NO									NO															NO		
z																													NO		
f	NO		NO							NO																				NO	
z																															
ɕ	NO																														
ʂ																															
x			NO																												
n			YES	NO				NO		NO					NO				NO											NO	
ɲ																															
m	YES																														
ɲ																															
l			YES	YES	YES			YES		YES							YES	YES								NO	NO				
r	YES		YES	YES	YES			YES		YES			YES	YES			YES	YES							NO	NO			NO		
w	YES		YES	YES	YES			YES		YES			YES	YES			YES	YES							YES	YES			NO		
j			YES	YES	YES					YES							YES	YES							YES	YES					
ɰ	YES													YES	YES										YES	YES					
ɰ																															

Figure 1. A dispersion matrix for initial CC clusters according to NAD.

	p	b	t	d	k	g	ts	tʃ	dʒ	tʃ	dʒ	f	v	s	z	ʃ	ʒ	ɕ	ʂ	x	n	ɲ	m	ŋ	l	r	w	j	w	j	
p				NO											NO			NO			NO		YES			YES	YES	YES	YES		
b																NO							YES		YES	YES	YES	YES	YES		
t						NO							NO					NO				YES	YES	YES	YES	YES	YES	YES			
d			NO											NO				NO				NO	NO	YES		YES	YES	YES	YES		
k				NO									NO						NO				NO	YES	YES	YES	YES	YES	YES		
g					NO								NO	NO					NO					NO	YES	YES	YES	YES	YES		
ts																								YES		YES	YES	YES	YES		
tʃ			NO										NO	NO												YES					
dʒ									NO		NO																	YES			
tʃ										NO																			YES		
dʒ												NO												YES							
f	NO			NO				NO		NO					NO		NO		NO		NO					YES	NO	YES	YES	YES	
v					NO							NO				NO		NO				NO	NO	NO	NO	YES	YES	YES	YES	YES	
s	NO			NO		NO		NO					NO			NO		NO				NO	NO	NO	YES	YES	YES	YES	YES		
z		NO		NO		NO							NO				NO		NO				NO	NO	NO	YES	YES	YES	YES		
ʃ					NO								NO											NO	YES	YES	YES	YES			
ʒ		NO											NO											NO	NO	YES	YES	YES	YES		
ɕ													NO											NO	YES	YES	YES	YES			
ʂ													NO											NO	YES	YES	YES	YES			
x								NO				NO	NO				NO								YES	YES	YES	YES			
n																															
ɲ																								NO	NO		NO	YES	YES	YES	
l																															
r									NO					NO																	
w			NO																												
j																															
w																															
i																															

Figure 2. A dispersion matrix for final CC clusters according to NAD.

	p	b	t	d	k	g	ts	dz	tʃ	dʒ	tʃ	dʒ	f	v	s	z	ʃ	ʒ	ɕ	ʂ	x	n	ɲ	m	ŋ	l	r	w	j	ɰ	j	
p																																
b			YES	YES			YES	YES		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
t		YES		YES		YES		YES		YES		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	NO	YES	NO	NO	NO	NO	NO	NO	
d		YES	YES		YES		YES			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	NO	NO	NO	NO	NO	NO	
k		YES	YES	YES			YES			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	NO	NO	NO	NO	NO	NO	NO	
g			YES	YES			YES			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	NO	NO	NO	NO	NO	NO	NO	
ts			YES	YES		YES				YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	NO	NO	NO	NO	NO	NO	NO	
dz				YES										YES								YES	YES	YES	YES	NO	NO	NO	NO	NO	NO	
tʃ		YES	YES	YES		YES		YES		YES							YES					YES	YES	NO	NO							
dʒ		YES	YES	YES		YES				YES												YES	YES	NO	NO							
tʃ																																
dʒ																																
f			YES	YES		YES		YES		YES							YES		YES		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
v			YES	YES		YES		YES		YES							YES		YES		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
s		YES		YES		YES		YES		YES				YES	YES							YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
z		YES		YES		YES		YES		YES				YES	YES							YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
ʃ		YES		YES		YES		YES		YES				YES	YES							YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
ʒ		YES		YES						YES				YES	YES							YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
ɕ		YES												YES	YES							YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
ʂ		YES												YES	YES							YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
x			YES	YES		YES		YES		YES			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
n			YES	YES		YES		YES		YES			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
ɲ			YES	YES		YES		YES		YES			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
m		YES	YES	NO		NO		NO		NO			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	
ŋ																																
l			NO	NO	YES	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES	NO	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	
r		NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	
w		NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES	
j		NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES	YES	

Figure 3. A dispersion matrix for medial CC clusters according to NAD.

adequate visualisation of the nature of NAD predictions. The matrices have been derived for the CC pairs in all word-contexts (initial, medial, final) for 31 Polish consonants (ordered by manner/place in columns/rows). They demonstrate (a) the non-overlapping sets of clusters in initial, medial and final position (which testifies to the logical integrity of the preference formulae: a given cluster cannot be equally

good in different word positions), (b) the distinct sets of preferred and dispreferred clusters (YES's and NO's), and (c) the continua of preference across the matrices from the most to the least preferred cluster (e.g., in Figure 1, *pj* is the most preferred cluster, followed by *pw* and *bj* (with the same value of NAD), followed by *pr* and *bw* (ditto), followed by *pl*, *br*, *tw* and *dj* (ditto), etc.). The key attribute of the matrices is that they reflect simultaneously the dichotomic (categorical) and scalar (gradual) evaluation of clusters. Human perception is both categorical and gradual, too.

4. NAD and the data

4.1. NAD and frequency

The consonant sequences were extracted from three types of corpora of Polish: a list of lemmas based on a 8 thousand dictionary of 'core' Polish ('dictionary' henceforth), a list of inflectional forms generated on the basis of standard paradigms for Polish nouns, adjectives and verbs ('paradigms' henceforth), and a list of over 500,000 inflectional forms derived from a collection of newspaper texts ('corpus', henceforth). The result was a list of approximately 1,500 consonant cluster types of all sizes and in all positions.

Nearly half of the NAD-preferred 60 initial CC clusters from the Polish matrix (cf. Figure 1 above) were compared to first 5 most frequent Polish initial CC clusters in the 'corpus', the 'paradigms' and the 'dictionary' and first 14 most frequent Polish initial CC clusters in the 'dictionary' and the 'corpus'. We noted a difference between 'corpus' and 'dictionary' in types of clusters and a difference between the sets of 5 and 14 clusters in the percentage of unmarked clusters. Table 3 below summarizes the results. Most importantly, there is an essential, qualitative difference in what we conclude depending on how much of the data we consider. If we look arbitrarily at the first 5 most frequent clusters in the language, 3 out of 5 are marked, which points to NAD as a weak predictor of frequency of occurrence. If, however, we examine a larger set (in this case 14 most frequent clusters were considered due to a notable drop in frequency of further clusters), it turns out that 50% of the clusters in the 'paradigms' and 'corpus' and 71% of the clusters in the 'dictionary' are NAD-preferred, i.e. relatively unmarked. Therefore, NAD scores as a good predictor for the 'dictionary' data only.

Table 3. Initial CC clusters: NAD and frequencies.

preferred according to NAD									frequent in 'paradigms' & 'corpus'	frequent in 'dictionary'
5	4.25	4	3.5	3.3	3.25- 3.2	3	2.5	2.3-2	pj, pr, st mj, sp	pj, pr, st, sp, pj
<u>pj</u>	kw <i>pw</i>	bj	fj <i>kj</i>	<u>pr</u>	bw gw <i>kr</i> tw xw	dj	gj <i>kl</i> tjfw tsw vj	<i>tr</i> sw <i>gr</i> <u>mj</u> pl zj	dl, vj, kt. gd, <i>pj</i> , <i>kj</i> . zn, <i>kr</i> , <i>tr</i>	<i>vj</i> , <i>tr</i> , <i>kr</i> , sw, <i>gr</i> , <i>sk</i> , <i>pw</i> , <i>kl</i> , <i>mj</i>
<u>preferred according to NAD in 5 most frequent clusters (underline)</u>										
<i>preferred according to NAD in 14 most frequent clusters (italic)</i>										

Next, if we examine the marked clusters from among the most frequent 14 in the 'corpus', we notice a huge discrepancy between the 'dictionary' and the 'corpus' frequencies (roughly reflecting the difference between type and token frequencies, respectively; cf. Table 4). This is accounted for by usage: the functional load of words containing many of these marked clusters is extremely high in Polish.

Table 4. Initial CC clusters: type vs. token frequency.

cluster	dictionary	corpus	words
/dl /	3	172698	<i>dla</i> 'for'
/kt/	7	92997	<i>kto</i> 'who', <i>który</i> 'which'
/gd/	3	75041	<i>gdy</i> 'when'
/zn/	19	35424	<i>znać</i> 'know'

Earlier in Table 3 we saw the failure of NAD to predict the frequent clusters in the 'corpus' (50% score). Data in Table 4 shows that 4 out of 7 dispreferred clusters (i.e. another 29% of all 14) have very low type frequency, which gives credit to NAD that predicts them to be dispreferred. In other words, in this case NAD relates well with type frequency and not at all with token frequency.

4.2. NAD: criteria manipulation

NAD calculations may show different results depending on the choice and use of criteria. We have already mentioned scalar vs. binary qualification of the 'goodness' of clusters (cf. section 3.4. above). The matrices (Figures 1-3 above) show both binary and scalar typology

of clusters. The analysis of frequencies demonstrates the advantage of scalar qualification since it reflects *relative* markedness, predicted by the *preference* theory (cf. Dressler 1999). For instance, in Table 3 above the “preferred according to NAD” clusters are distributed along 9 values within the category *preferred*.

The set of features used for the calculations may be changed or expanded. For example, laryngeal features are non-redundant within obstruents and largely redundant within sonorants. Therefore, instead of Lx (difference in the value of voicing),⁵ the S/O criterion, i.e. the difference between sonorant and obstruent (set at 1) has been introduced into the equation (cf. section 3.2 above). It makes a difference for (calculation formulae as in section 3.3 above):

–initial clusters of voiced obstruents with sonorants (in this order; they improve), e.g.

/vj/ 2.5 (Lx contributes 0) 3.5 (S/O contributes 1)

/gr/ 2.2 (Lx contributes 0) 3.2 (S/O contributes 1)

–initial clusters of voiced obstruents (they worsen), e.g.

/zd/ -4 (Lx contributes 0) -5 (S/O contributes |-1|)

More detailed features of MOA and POA could be introduced and their values could be modified, for example by introducing weights (e.g. Calderone & Bertinetto 2013 experimented with the proposal of 1.0 for CV opposition, 0.8 for manners and 0.5 for places and voicing as input values to their probabilistic system). Language-specific values can be introduced (cf. NAD for Polish vs. English) in order to reflect different organization of segmental inventories across languages, entailing different degrees of exploitation of a given feature. POA’s for vowels need to be included, too. Additionally, more phonetic detail could be considered when specifying the positional preferences, e.g. the timing differences between initial and final clusters.⁶

All the modifications aim at setting the values so that the calculations maximally match the data. In this way the model (deductive) receives feedback (inductive) from the data.

4.3. NAD and acquisition: collected data

The order of acquisition of different types of clusters appears to vary across languages. For example, Fikkert & Freitas (2004) showed the following sequences of acquisition of clusters by Dutch and European Portuguese children:

PL > FL > SP	Dutch children
SP > PL > FL	European Portuguese children

Thus, SP (sibilant-plosive) clusters were acquired last by the Dutch children and first by the European Portuguese kids. Additionally, some Dutch children had SP first. Frequency effect was ruled out since in both languages CL (obstruent-liquid) clusters (PL plosive-liquid and FL fricative-liquid) are much more frequent than SP (10% vs. 3% of the kids intake). Also, CL emerged long before SP (and SP long before CL respectively) in both languages (at the 3-6 months span). As we will see, for Polish children none of this is true.

4.4. NAD and acquisition: experimental data

A picture naming task was used to elicit the production of words with six initial clusters /st, sp, sk, sx, sm, sw/ by 49 Polish kindergarten children, half of them with normal phonological development, mean age of 42 months, the other half with DPD (developmental phonological disorder), mean age of 57 months. The children were selected out of 114 according to the results of a battery of tests testing their overall performance, among others their ability to produce singleton consonants included in the studied clusters (Yavaş and Marecka 2013). The extent of reductions from the most reduced to the best preserved was as follows:

$$sx > sw > sk > sp > sm > st$$

As can be seen, both ends of the scale as well as the middle are occupied by marked (=dispreferred) clusters (signalled by italics). The group of children with DPD reduced the majority of the productions, but showed similar tendencies (cf. also Marecka & Dziubalska-Kołaczyk 2014).

However, the experiments involved an arbitrary selection of the data for elicitation, without reference to the systems of the languages studied. For instance, for word initial *st* in Polish an input frequency effect may be considered, since this is the third most frequent cluster in Polish (both type and token frequency). As Fikkert & Freitas (2004: 10) observed “it is important to consider the language system as a whole to interpret the data, both to explain differences between children acquiring the same language (i.e. the child’s own phonological system determines what optimal realizations for clusters are), and between children acquiring different languages”.

4.5. NAD, frequency and acquisition

Table 5 presents once again the data of Table 3, but with the addition of data coming from one case study, namely, the analysis

of clusters produced by Zosia, a young girl from Poznan, who was recorded between 1;7 and 3;2 years of age according to a longitudinal production study (cf. Zydorowicz 2010). Marked clusters that are frequent in Zosia's speech (shown as strikethrough) can be explained by a strong frequency effect, i.e. by statistical learning. These acquisitional data furthermore confirm that small sets of data (e.g. 5 early clusters) are not sufficient to draw conclusions about the predictive power of NAD, whereas a larger pool of data demonstrates that:

- the majority of frequent clusters is unmarked in the 'dictionary' (78%)
- approximately 50% of frequent clusters are unmarked in the 'corpus'
- approximately 30% of first acquired clusters are unmarked.

The top cluster (both as type and token frequency) is the same in the most frequent list and in the first acquired list; this cluster is /pʃ/, and it is a marked structure according to NAD. Its occurrence can be explained by morphological productivity since /pʃ/ appears within highly productive affixes, e.g. in *przed-*, *przy-*, *prze-*, all beginning with /pʃ/.

Table 5. Initial CC clusters: NAD, frequencies and L1 acquisition.

preferred according to NAD									Frequent in ‘paradigm’ & ‘corpus’	Frequent in ‘dictionary’	L1 acquisition by Zosia 2.8-2.9
5	4.2 -5	4	3.5	3.3	3.25 -3.2	3	2.5	2.3 -2	pf. pr, st. mj. sp dl, vj. kt. gd, pj, kj. zn, kr. tr	pf. pr, st. sp- pj vj. tr. kr. sw, gr. sk, pw, kl. mj	(33) pj, (12) st. (9) kc, (7) mj, (6) cp, (5) dv. kl, xts, (4) sp, tjt, sw, xw, (3) pw. gw, sk, mn
<u>pj</u>	kw pw	<u>bj</u>	<u>fj</u> kj	<u>pr</u>	bw gw kr tw xw	dj	gj kl tjw tsw vj	tr sw gr <u>mj</u> pl zj			
preferred according to NAD in 5 most frequent clusters (<u>underline</u>) <i>preferred according to</i> NAD in 14 most frequent clusters and L1 (<i>italic</i>) overlapping marked in frequent clusters and L1 (strikethrough) The numbers in parentheses in the last column of Table 5 are tokens in Zosia’s speech.											

4.6. NAD, cluster size and ease

In the above sections we discussed frequent and early clusters and how they score in terms of markedness. Let us consider 'easy clusters' in this perspective, too. Of the two measures of markedness of clusters used in this study, NAD and the size of a cluster, the latter is relatively straightforward with respect to ease: a shorter cluster is

easier to articulate and to hear than a longer one. NAD's predictions of ease are no longer so straightforward since NAD is a derivative of articulatory and acoustic features contributing in total to the final auditory effect. Characterising ease from this perspective is ambiguous. We do not know, for instance, what is easier: sustaining a gesture or changing (=modulation of) the gesture (cf. e.g. /wu/ or /ji/ in which a labial or palatal gesture is sustained while the perception of those clusters is very poor). There is an apparent conflict between perceptual contrast and ease of articulation; however, the auditory distance obtained within a cluster is the result of an optimal compromise between the two. In casual speech, clusters reduce and change, and the modifications are much more often due to assimilation than dissimilation (i.e., due to the tendency towards ease of articulation). Still, although the general tendency is to reduce clusters, marked (=dis-preferred) clusters arise, too (cf. Dressler *et al.* 2001), so the tendency towards ease may result in the actual markedness.

To provide an illustration of the relationship between markedness and ease, let's consider long clusters, first. We classified 4, 5 and 6 consonant clusters as long (cf. Table 6). We analysed 273 long clusters, which constituted 15% of all clusters of Polish (cf. section 4.1. above). The low percentage of the types (for example, only 3 clusters of 6 consonants) implies that the size of a cluster does predict difficulty of production and, consequently, of perception (i.e., long clusters are difficult and thus rare). Additionally, almost 100% of those clusters are morphonotactic (see Table 7). Morphonotactic clusters are expected to be marked since in this way they signal morphological boundaries (cf. Dressler & Dziubalska-Kołaczyk 2006). In the case of morphonotactic long clusters, however, there is the additional aspect of size resulting from the concatenation of morphemes that already contain clusters. Such sizeable clusters are bound to be difficult to pronounce. Markedness expressed by size is thus a good predictor of phonetic difficulty .

Table 6. Cluster length in Polish.

number of C	number of types
2	485
3	976
4	219
5	51
6	3
	1734

Table 7. Long clusters: phonotactics vs. morphonotactics.

Cluster length (in number of phonemes)	Lexical (intra-morphemic)	Morphono-tactic	Examples
6	–	100%	prze <u>st</u> ęp <u>st</u> wie, wewn <u>ą</u> tr <u>z</u> współnotowy, wewn <u>ą</u> tr <u>z</u> związkowy
5	–	100%	dzieci <u>ń</u> stwie, bez <u>w</u> zględn <u>ie</u> , roz <u>br</u> zmiewa, kontr <u>pr</u> zyk <u>ł</u> ad, post <u>z</u> wiązkowy, tysią <u>c</u> stronicowy
4	5%	95%	/rstf/ war <u>st</u> wa (lex 3%) minister <u>st</u> wo (-stwo morph 97%), <u>kr</u> wi <u>ą</u> (lex), krn <u>ą</u> brny (lex), <u>ż</u> d <u>z</u> ęb <u>ł</u> o (lex), roz <u>t</u> kliwi <u>ą</u> c (morph), odstr <u>z</u> ął (morph), j <u>ę</u> dny (lex)

Markedness measured by NAD, however, may fail as a predictor of difficulty in specific types of performance, in particular, when marked (NAD dispreferred) clusters arise in contexts triggered by the principle of least effort. The application of phonostylistic processes in casual speech, such as vowel reduction “may lead to a temporary rise of ‘new’ clusters which are marked” (Dressler *et al.* 2001: 108). In the casual, everyday style, we found the following ‘new’ clusters in Polish (cf. Dressler *et al.* 2001: 108, Table 8), all of them marked. In this case, the natural speaker-friendly tendency for ease led speakers to produce long clusters.

Table 8. “New” clusters resulting from vowel reduction.

Double clusters (12)	nč	ŋ č	ʒč	m̩ m	ʃ s	ss(2)	m̩ v	zdz	zz	ʒ ts	ŋ d
Triple clusters (9)	tlk	t ɫ k	pɫk	fʃs	m̩ vj	pʃn	ssn	pʃp	ʃ td		
Four-consonant clusters (4)	fʃšk	pkf̩w	pɫf̩w (2)								
Five-consonant clusters (2)	fʃstk	fʃsts									

5. Conclusions

We have shown a better implicational relation between the unmarked and frequent clusters than between the unmarked and first acquired ones. Additionally, unmarked implies frequent in a dictionary better than in a corpus. Generally, markedness predictions based on NAD appear to work better for the system (lexicon) and its use (dictionary and corpus) than for its acquisition. It may be that we deal with a critical time in first language acquisition when the

markedness predictions begin to operate. This, however, remains to be investigated. Markedness predictions of ease of production have been only shortly discussed in the paper and shown to depend on the applied measure of markedness (size of cluster vs. NAD) as well as the type of performance.

We conclude that for markedness to be a reliable tool of analysis we need to consider the whole system of the language studied, cross-compare to other systems, discover language-specific constraints, collect and elicit larger quantities of reliable, natural, representative and detailed data, as well as discover more criteria which co-determine markedness.

Address of the Author

Katarzyna Dziubalska-Kołaczyk Faculty of English, Adam Mickiewicz University in Poznań, al. Niepodległości 4, 61-874 Poznań, Poland <dka-sia@wa.amu.edu.pl>

Notes

¹ Michał Jankowski, Paulina Zydorowicz.

² Morphotactics is the area of interaction between morphotactics and phonotactics (cf. Dressler and Dziubalska-Kołaczyk 2006).

³ As opposed to a purely phonemic treatment of clusters, reflected in their phonemic transcription.

⁴ For the purposes of the calculations, the phonotactic calculator has been developed (consecutive versions by Krynicki, Pietrala, the current one by Aperliński), available online under <http://wa.amu.edu.pl/nadcalc/>.

⁵ We included the voicing contrast, symbolized by Lx, in the previous version of NAD. The presence of the voicing contrast between two members of a cluster was signaled by the value 1, its absence by 0.

⁶ "Onset clusters are considered to be timed more rigidly, allowing for less timing variability and being less overlapped than heterosyllabic or coda clusters (Byrd 1996; Hardcastle & Roach 1997)" (Bombien 2011: 35).

Bibliographical References

- Baroni Antonio 2014. On the importance of being noticed: the role of acoustic salience in phonotactics (and casual speech). *Language Sciences* 46. 18-36.
- Bombien Lasse 2011. *Segmental and prosodic aspects in the production of consonant clusters – On the goodness of clusters*. PhD dissertation, München.

- Calderone Basilio & Pier Marco Bertinetto 2013. From Phonotactics to Syllables. A psycho-computational approach. Talk delivered during the 46th *Societas Linguistica Europaea*, 17-20 Sept, 2013, Split.
- Dressler Wolfgang U. 1999. On a semiotic theory of preferences in language. *The Peirce Seminar Papers* vol. 4. New York: Bergham Books. 389-415.
- Dressler Wolfgang U., Dziubalska-Kolaczyk Katarzyna & Rossella Spina 2001. Sources of markedness in language structures. *Folia Linguistica Historica* 22 (1-2). 103-135.
- Dressler Wolfgang U. & Katarzyna Dziubalska-Kolaczyk 2006. Proposing morphonotactics. *Rivista di Linguistica* 18(2). 249-266.
- Dziubalska-Kořaczyk Katarzyna 1999. Early L1 clusters and how they relate to universal phonotactic constraints. *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Francisco. 317-320.
- Dziubalska-Kořaczyk Katarzyna 2002. *Beats-and-Binding Phonology*. Frankfurt/Main: Peter Lang.
- Dziubalska-Kořaczyk Katarzyna 2009. NP extension: B&B phonotactics. *PSiCL* 45 (1). 55-71.
- Dziubalska-Kořaczyk Katarzyna 2014. Explaining phonotactics using NAD. *Language Sciences* 46. 6-17.
- Fikkert Paula & Maria Jořo Freitas 2004. The role of language-specific phonotactics in the acquisition of onset clusters. In L. Cornips & J. Doetjes (eds.), *Linguistics in the Netherlands 2004*. Amsterdam: John Benjamins. 1-12.
- Hyman Larry 2008. Universals in phonology. *The Linguistic Review* 25 (1-2). 83-137.
- Ladefoged Peter 2006. *A Course in Phonetics*. 5th edition. Boston: Heinle & Heinle.
- Marecka Marta & Katarzyna Dziubalska-Kořaczyk 2014. Evaluating the models of phonotactic constraints on the basis of the sC cluster acquisition data. *Language Sciences* 46. 37-47.
- Ohala John J. 1990. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. The Parasession on the Syllable in Phonetics and Phonology. *CLS* 26,2. 319-338.
- Olender Adam 2013. *A synchronic investigation of leftmost /s/ + stops*. M.A. dissertation, Adam Mickiewicz University in Poznań.
- Pietrala Dawid 2014. *The application of linguistic principles in programming: phonemic transcription and phonotactic constraints*. Unpublished Ph.D. dissertation, Adam Mickiewicz University in Poznań.
- Yavař Mehmet & Marta Marecka 2013. Acquisition of Polish #sC clusters in typically-developing children and in children with phonological disorders. *International Journal of Speech-Language Pathology*.
- Zydorowicz Paulina 2010. Consonant clusters across morpheme boundaries: Polish morphonotactic inventory and its acquisition. *PSiCL* 46(4). 565-588.
- Zydorowicz Paulina, Orzechowska Paula, Jankowski Michał, Dziubalska-Kořaczyk Katarzyna, Wierchoń Piotr & Dawid Pietrala In press. Phonotactics and morphonotactics of Polish and English Theory, description, tools and applications. Poznań: Adam Mickiewicz University Press.

Phonotactic vs. morphonotactic obstruent clusters in Slovak and German

Wolfgang U. Dressler, Miroslava Hliničanová, Matej Ďurčo, Karlheinz Mörrh, Katharina Korecky-Kröll

Slovak has more consonants than German and also more consonant clusters (plus accented syllabic sonorants). In addition, their asymmetric distribution within the word is different: Slovak has more clusters in word-initial position, German many more in word-final position. This difference is reinforced by morphology, insofar as only Slovak has monoconsonantal prefixes, and only German has monoconsonantal suffixes. The main word-internal difference is that in compounding German increases clusters, whereas Slovak decreases them. Asymmetries in terms of type frequency are radicalised in token frequency, i.e. in the profitability of clusters. In addition, clusters arise in Slovak also via vowel deletion, which is extremely rare in German (except in casual speech and in dialects).

1. Introduction

In continuation of previous theoretical and contrastive work (Dressler & Dziubalska-Kořaczyk 2006, Dressler *et al.* 2010, Korecky-Kröll *et al.* 2014) we contrast Slovak and German patterns of consonantal morphonotactics vs. phonotactics from a phonological, morphological and corpus-linguistic perspective. So far we have done only some corpus-linguistic work on German, but the ongoing PhD thesis of the second author allowed a vast corpus-linguistic study of Slovak as well (which has never been done before). This enables us to discuss it also in the framework of typological differences between a Slavic and a Germanic language.

The theoretical background is Natural Phonology and Morphology (cf. Dziubalska-Kořaczyk & Weckwerth 2002, Kilani-Schoch & Dressler 2005), and this approach does not only strive towards descriptive und explanatory adequacy but also towards guaranteeing, at least partially, the psychological reality of the linguistic constructs. This demands a psycholinguistic perspective. In contrast to Korecky-Kröll *et al.* (2014), the psycholinguistic perspective of this contribution is devoted to acquisition and not to processing. In usage-based linguistic and psycholinguistic approaches (Bybee 2001, Bauer 2001, Tomasello 2003) it is often claimed that token frequency is important only for the question of storage (which is not an issue

here), whereas only type frequency and discrepancy between high type frequency and low token frequency is relevant for productivity and profitability of patterns (cf. Du & Zhang 2010, Berg 2014). We compared type and token frequency, in order to evaluate these claims with fresh data.

We investigate prototypical and non-prototypical cases of morphotactics, i.e. the prototypical case of merely concatenative shapes of morpheme combinations, particularly when they differ from the phonotactics of lexical roots and morphemes and thus signal morpheme boundaries, as in English *seem-ed* /si:m-d/ (i.e., there is no lexical final [-md] cluster in English). All German cases that we investigate are of this type. Slovak has, in addition, also the non-prototypical case of morphological combinations resulting in vowel deletion, which is extremely rare in German, for example in *Risiko* 'risk', adj. *risk-ant* 'risky'. Slovak has many more, e.g. *pes* 'dog', pl. *ps-y*, adj. *ps-í*.

In those cases in which the pattern investigated occurs both in lexical phonotactics and morphophonotactically, such as word-final [pst] in German *Papst* 'pope' vs. *du gab-st* 'you gave', we will look also at the quantitative distribution of both patterns in large electronic corpora. In order to focus on the most radical contrasts between German and Slovak, we limit our study to obstruent clusters.

Consonant clusters are studied within the framework of Dziubalska-Kořaczyk's (2002) Beats-and-Binding model (cf. Dziubalska-Kořaczyk & Zielińska 2010), which starts with syllabic nuclei as beats and consonants bound to them, but does not assume syllabic boundaries. For the sake of better understanding we will however provide information about traditional syllabification for each group of examples. Words are defined as morphological words and thus include compounds (which also have just one main accent).

The corpus-linguistic research is done with the help of the huge electronic corpora of the Slovenský národný korpus (Slovak National Corpus), which contains 829 million word tokens, the Austria Media Corpus, which contains all the content of all Austrian journals and periodicals of a period comprised between 15 and 5 years ago (7 billion word tokens), and the Austrian Academy Corpus, which focuses on written texts of all genres of the nineteenth and earlier twentieth centuries (half a billion word tokens).

Examples are written in the national orthographies. In the German consonantal system, *ch* is a voiceless palatal or velar fricative, *sch* (and word-initial *s* before a stop) a voiceless shibilant. Its Slovak counterpart is written *š*, its voiced correspondent *ž*, *č* stands for a voiceless palatal affricate, *dž* for its voiced counterpart, *z* is the

voiced counterpart of *s*; *c* (as in German) and *dz* the corresponding affricates, *ch* a voiceless prevelar fricative, *h* its voiced counterpart, *ť* and *ň* are the palatalized counterparts of *t* and *n*; an accent above a vowel grapheme means length; *r*, *l* between obstruents are syllabic. In Slovak there is a biunique relation between phonemes and graphemes, with the exception of word-final obstruent devoicing and forward devoicing of voiced to following voiceless obstruents (both processes also occur in German, as well as the digraph *ch*).

2. Word-initial position

Word-initially, Slovak is much richer in obstruent clusters than German, i.e. due to morphonotactic concatenation through the prefixes *z-*, *s-*, *v-*, e.g. in *s+chrad+nú-t'* 'to get older', *z-drav-i-t'* 'to greet', *v-kvap-k-a-t'* 'to drip in' and because of greater phonotactic complexity, e.g. *pstruh* 'trout', *škvár-a* 'slag', *vdov-a* 'widow'. Since Slovak is very rich in word-initial obstruent clusters (which is not the case in German), Slovak has many obstruent clusters which are either purely phonotactic or purely morphonotactic or are ambiguous, i.e. both phonotactic and morphonotactic.

Purely phonotactic obstruent clusters are formed by the obstruents *džb* in *džbán* (2130 tokens) 'jug', *pch* in *pchat'* (1850) 'mend', more complex *pstr* in *pstruh* (3552) 'trout', *pšt* in *pštros* (920) 'ostrich'. Purely phonotactic clusters are also combinations with the affricate *č*: *čk* in *čkat'* (209) 'gulp', *čp* in *čpavok* (661) 'ammonia' and *čv* in *čvachtat'* (124) 'clap', more complex *škv* in *škvarka* (313) 'bacon greave'. Of course, one has to classify the syllabic sonorants /*r*, /*l* as non-consonantal, as in *štrngnúť* (520) 'impinge', *žbrnda* (220) 'sword'.

Purely morphonotactic obstruent clusters are *s-č* in *s-černiet'* (351) 'to become black' or *s-t'* in *s-táhovat'* (14463) 'drag along' or *v-b* in *v-biehat'* (244) 'come in', *v-bodnúť* (93) 'prick', *v-bit'* (48) 'strike in'.

Slovak ambiguous obstruent clusters, which are both phonotactic and morphonotactic, are restricted to combinations with word-initial obstruents *s*, *z*, *v* (only exception: *ps*, see below), e.g. *sc* in *scenár* (31189) 'scenario' vs. *s-c* in *s-cvrknúť* (1475) 'shrink', *sch* [sx] in *schizma* (494) 'schisma' vs. *s-ch* in *s-chladit'* (2229) 'chill', *sf* in *sféra* (30612) 'sphere' vs. *s-f* in *s-formovat'* (3628) 'form', *sk* in *skala* (25427) 'rock' vs. *s-kombinovat'* (1576) 'combine', *vč* [fč] in *včera* (201703) 'yesterday' and *včela* (8429) 'bee' vs. *v-č* in *v-čas* (14111) 'in time', *v-člen-it'* (909) 'intersperse', *vzl* in *vzlyk* (1250) 'sob' vs. *v-zl* in *vzlet* (868) 'ascent', *zb* in *zbojník* (5799) 'robber' vs. *z-b* in *z-balit'* (4266) 'pack up', *zbr* in *zbraň* (78918) 'weapon' vs. *z-br* in

z-brúsiť ‘smooth’, *zv* in *zvierá* (63895) ‘animal’ vs. *z-volit’* (43706) ‘choose’. The exceptional cluster *ps-* is phonotactic in *psi* ‘slag’, morphonotactic as result of morphology-induced vowel deletion of the root vowel of *pes* ‘dog’ in Pl. *ps-y* and in the derived adjective *ps-í* (7916), *ps-in-ec* (98) ‘dog den’, compound *ps-o-vod* (2443) ‘dog guide’.

Table 1 additionally provides information about the proportion of clusters that can be both phonotactic and morphonotactic in the lexicon. The as phonotactic clusters include the “ambiguous” clusters as well; the same holds for the morphonotactic counterparts.

Table 1. Slovak word-initial obstruent clusters

CLUSTERS	TOKENS	LEMMAS
morphonotactic (M)	42 058	58
phonotactic (P)	12 073 799	1 884
ambiguous / thereof morphonotactic	2 360 612	930
ambigues / thereof phonotactic	23 177 076	2 754
Total	37 653 545	5 626

Clearly word-initial phonotactic clusters (including those occurring in inflected word-forms) largely outweigh morphonotactic types and this contrast is radicalised in tokens. We interpret this as showing that the relative normalcy (cf. Wurzel & Schentke 1989) of phonotactic sequences within a language is not only a question of type frequency (i.e. number of lemmas or inflectional word forms in which each sequence occurs), but also a question of their profitability in term of tokens. This is especially important for consonant clusters, because very frequent words are known for their tendency to reduce consonant clusters.

The analysis shows that most clusters are of the ambiguous group and that the marked group of purely morphonotactic clusters is extremely limited. In other words, morphonotactic clusters are nearly always built on the patterns of existing phonotactic clusters.

The German standard has no monoconsonantal prefixes, in contrast to Bavarian-Austrian dialects, as in *g’storben* ‘died’, *b’soffen* ‘drunk’, *z’ruck* ‘back(wards)’, etc., corresponding to Standard German *ge-storb-en*, *be-soff-en*, *zu(-)rück*. And the in the highest level of phonotactic complexity is reached in the sequence of two obstruents plus a sonorant, as in *Strafe* ‘fine’, *Sprache* ‘language’, *Splitter* ‘splinter’, *Strontium*, *Skrupel* ‘scruple’, *Sklave* ‘slave’.

The other German word-initial obstruent clusters are: *[kv]* as in

Quelle ‘source’, [ks] as in *Xenophobie*, [ps] as in *Psychologie* ‘psychology’, *zw* [tsv] as in *Zwang* ‘coercion’, *sc*, *sk* [sk] in loanwords such as *scannen*, *Skalp*, *st* in loanwords such as *Stil* ‘style’ (compare *st* [ʃt] in *Stock* ‘stick’), *sp* in loanwords such as *Spatium* ‘space’ (vs. *sp* [ʃp] in *Spaß*), [sf] in loanwords such as *Sphäre* ‘sphere’, [sv] in loanwords such as *Sweater*, [sç] in loan words such as *Schizophrenie* ‘schizophrenia’, *schw* [ʃv] as in *Schwanz* ‘tail’; limited to learned loanwords, one should also mention *pt*, *ft*, such as *Pteridin*, *Phthisis*, *cht* [xt], such as *chthonisch*, *sz* [sts] as in *Szene* ‘scene’.

3. Word-internal position

Word-medially both languages have comparable phonotactic consonant clusters. In German, the syllable boundary is in compounds immediately before the second compound member and in prefixations always after the prefix, otherwise word-internal syllable onsets always follow the pattern of word-initial onsets. In Slovak, the syllable boundary is always after the first consonant of clusters (Dvonč *et al.* 2013).

Morphonotactic obstruent clusters arise in word formation in different distributions according to the typological difference between Germanic compounding languages and the rather derivational Slavic languages. Thus German compounds such as *Dienst+grad* ‘military rank’ have no correspondence in Slovak, not only because common Slovak is much poorer in compounds than German, but also because for Slovak the compounding interfix *-o-* is typical, as in *štrk-o-piesky* ‘gravel beach’ (from *štrk* ‘gravel’, *piesok* ‘sand’, Pl. *piesk-y*), *mozg-ov-o-ciev-n-y* ‘cerebrovascular’ (*mozog* ‘brain’, Adj. *mozg-ov-ý*, *ciev-a* ‘vessel’, Adj. *ciev-n-y*). In contrast, German has the obstruent interfix *-s-*, which increases consonantal complexity, as in *König-s-krone* ‘royal crown’, *Fried-hof-s-mauer* ‘cemetery wall’.

Slovak word formation creates, via many consonant-final prefixes and consonant-initial suffixes, many complex morphonotactic obstruent clusters, such as *roz-drob-en-ý* ‘broken up’, *gróf-stvo* ‘county’. Furthermore, in contrast to German, Slovak declension creates via vowel deletion new morphonotactic obstruent clusters, as in *mozog* ‘brain’, Gen.Sg. *mozg-u* vs. *vosk* ‘wax’, *laket* ‘ellbow’, Gen.Sg. *lakt’a* vs. *akt* ‘act’, *otec* ‘father’, Gen.Sg. *otc-a*, *list-ok* ‘leaflet’, Gen.Sg. *list-k-a / u*, *o-pas-ok* ‘belt’, Gen.Sg. *o-pask-a* vs. *po-tok* ‘creek’, Gen.Sg. *po-tok-a / u*, prefixed from *tok* ‘river, stream’.

Since Slovak does not allow word-final obstruent groups *-zg*, *-tk*, *-stk*, one could be induced to assume that we do not deal here with

vowel deletion, but with automatic phonological vowel insertion in the Nom.Sg. /mozg, list-k/, etc. As a consequence, these obstruent clusters would not be morphonotactic, but phonotactic. However with some obstruent clusters, these occur also word-finally, i.e. one would have to assume a non-automatic, lexically restricted morphonological vowel insertion, for example in Nom.Sg. *o-pas-ok* 'belt', Gen. *o-pas-k-u/a* (a morphosemantically opaque diminutive derivation from *pás* 'strap, belt') vs. *vosk* 'wax', *stisk* 'pressure', *vý-prask* 'clout', *kiosk* 'kiosk'. Note that the assumed underlying /mozg/ would be subject to obligatory word-final obstruent devoicing to [mosk]: this word-final obstruent cluster would be perfectly pronounceable.

Of the other obstruent clusters we would like to cite the special case of disyllabic *chrbát* 'back', Gen. *chrbt-a/u* vs. the Latinate loanwords *recept* 'receipt', *adept*, *koncept*, *Egypt*. The insertion of a long [a:] in /xrbt/ → *chrbát* would not only be isolated, but also phonologically totally implausible. The assumed underlying Nom.Sg. /xrbt/ would be subject to phonological voicing assimilation /xrpt/. The isolatedness of the deletion of long -á- in all other inflection forms of Masc. *chrbát*, plus in the derivations Fem. *chrbt-ic-a* 'back bone', Adj. *chrbt-ov-ý* has stimulated the formation of rare alternative forms without vowel deletion. The token frequencies in the Slovak National Corpus are, in the case of Gen.Sg., *chrbt-a* 6368, *chrbt-u* 172 vs. *chrbát-a* 68, *chrbát-u* 19. This token frequency distribution is further evidence for the basic status of /a:/ in *chrbát*.

Furthermore, vowel deletion (which is also assumed in Růžicka *et al.* 1966, Sokolová *et al.* 1999) instead of vowel insertion is also supported by the two following arguments. First, the alternating vowel is restricted to the unmarked base form of the Nom.Sg., whereas vowel deletion signals in an anticipatory manner (cf. Dressler 1985) marked inflection, derivation or compounding via the new morphonotactic consonant cluster; see e.g. the above-mentioned example *mozg-ov-o-ciev-n-y* 'cerebrovascular' (from *mozog* 'brain'). Second, the alternating vowels originated by sound change law from the Protoslavic ultra-short vowels *jer* and *jor*,¹ which later were automatically deleted in unstressed positions (cf. Krajčovič 1971, Sabol 1989, e.g. in Old Church Slavonic *xr̥bb̥t̥b̥* > Slovak, Serbo-Croatian *chrbát*, Bulgarian *chrbét*, Slovene *hrbét*). In current theories of language change there is no plausible explanation, why in these alternations vowel deletion should have been reanalysed as vowel insertion.

German correspondences with obstruent-initial suffixes are morphonotactic *Kind-chen* 'child-DIM', *Gast-lich-keit* 'hospitality' (/ç/ is the only obstruent followed by the allomorph -keit of -heit in German, thus obstruent clusters and especially plosive clusters are avoided),

lenk-bar ‘guidable’, *Schick-sal* ‘fate’, *furcht-sam* ‘timid’, *ab-wärts* ‘downwards’, unproductive *Jag-d* ‘hunting’, *Schlach-t* ‘battle’, *Schrif-t* ‘writing’, *Ein-kunf-t* ‘income’, *Ge-lüb-de* ‘vow’, *Klap-s* ‘smack’, *Knack-s* ‘crack’, *äch-zen* ‘to groan’, the isolated suffixoid in *Witz-bold* ‘witty fellow’.

Verb prefixation too creates new word-internal obstruent clusters. For example, the separable prefix/particle *ab-* motivates the exclusively morphonotactic clusters /p-d, p-t, p-g, p-k, p-ʃ, p-ts, p-v/, as in *ab-drehen*, *ab-treten*, *ab-geben*, *ab-kommen*, *ab-schaffen*, *ab-ziehen*, *ab-wickeln* (with the addition of longer clusters, as in *ab-streiten*). Also some of the few non-separable verbal prefixes create new clusters, as with *ent-*, and the earlier but now only vestigial suffix *ant-*. All nominal and adjectival prefixes are non-separable. In addition, prefixes (and compounding) create geminate consonants which are disallowed morpheme-internally, and, phonotactically even worse, pseudogeminates are created by syllable- and morpheme-final obstruent devoicing, as in *ab-bauen* with /p-ʃb/.

4. Word-final position

In word-final position, German has morphonotactic *-Cst* clusters in 2nd singular verb forms, superlatives and unproductive nominalisations with the homophonous *-st* suffix.

As Dressler & Dziubalska-Kołaczyk (2006) have shown, exclusive morphological motivation exists for the clusters /-xst, -fst/, as in *lach-st* ‘you laugh’, *schlaf-st* ‘you sleep’ as well as for the clusters /-pfst, -rkst -lkst, -nkst, -lpst, -mpst/.

A strong default is represented by the regional variant /-nkst/ (as in *häng-st* ‘you hang’ vs. monomorphemic *Angst* ‘anxiousness’ as well as by /-rpst/. I.e. there are extremely few phonotactic cluster occurrences vs. an extensive, potentially unlimited number of morphonotactic cluster occurrences. The default is slightly weaker (i.e. there are significantly more phonotactic exceptions) in postvocalic /-pst/, as in *lieb-st* ‘you love’ as well as in postvocalic /-kst/ and in the affricate /ts/ followed by /-(s)t/, as in *reiz-(s)t* ‘you/(s)he irritate(-s)’. And the morphonotactic default is even weaker with base-final sonorant as in /-rst, -nst/ clusters.

A further source of word-final morphonotactic obstruent groups is the nominal *-s* Gen.Sg., as in *des Lob-s* ‘of the praise’, *Prinzip-s* ‘principle’, *Kalb-s* ‘calf’, *Skalp-s* ‘scalp’, *Korb-s* ‘basket’, *Ge-zirp-s* ‘chirping’, *Schiff-s* ‘ship’, *Archiv-s* ‘archive’, *Schilf-s* ‘reed’, *Dorf-s*

'village', *Nerv-s* 'nerve', *Krieg-s* 'war', *Lack-s* 'varnish', *Werg-s* 'tow', *Quark-s* 'curd', *Talg-s* 'tallow', *Ulk-s* 'trick', *Dach-s* 'roof'.

Parallel phonotactic clusters occur in *Schnaps* 'brandy', *Gips* 'gypsum', *Raps* 'colza', *Kollaps* 'collapse', *Krebs* 'cancer', *Rülps* 'belch', *Knirps* 'dwarf', *Keks* 'cookie', *Koks* 'coke', *Klecks* 'blot', *Fuchs* 'fox'. Morphophonotactic word-final (*r/l*)*fs*, *xs*, *lks*, *rks* have no phonotactic parallels.

A problem is represented by imperatives of the type *knicks!*, *schubs!* 'curtsy!', 'push!'. First, it is unclear whether the word-final *-s* is synchronically still a derivational suffix. Second, even if not, it is unclear whether such imperatives are to be classified as base forms (if yes, then phonotactic) or as morphologically derived from the infinitive as lexical entry.

In contrast to German, Slovak word-finally has no consonantal inflection suffixes, but creates morphophonotactic obstruent clusters in suffixless genitive plurals via deletion of the stem vowel, as in *rod-i-sk-o* 'birth place', Gen.Pl. *rodísk*, *mzd-a* 'wage', Gen.Pl. *miezd*.

Examples of Slovak word-final obstruent clusters, which are both morphophonotactic and phonotactic, (with their token frequency values within brackets) are e.g. Gen.Pl. (all neuters) *prac-ov-í-sk* (9573) 'of the working places', *voj-sk* 'troups' (8015) vs. Nom.Sg. *zisk* 'yield' (20788) or *blesk* 'lightning' (4852); Gen.Pl. *účilíšť* 'educational establishments' (935) vs. *plášť* (2120) 'coat'; *miest* (9982721) 'towns' vs. *rast* 'growth' 37958. Among the female Gen.Pl., one can mention [st] in *hviezd* (153253) 'of stars' vs. *zjazd* (3239) 'assembly'; from *naft-a* 'petroleum', the Gen.Pl. *náft* (6) vs. *kšeft* 'shop' (from Austrian German *G(e)schäft*) (551); from *pošt+a* 'post', Gen.Pl. *pôšt* (4813) or from *vražď-a* 'murder', Gen.Pl. *vrážď* [vra:št] (4723) vs. *mušt* (214) 'must'; from *lopt-a* 'ball', Gen. Pl. *lôpt* (2819) vs. *koncept* 'concept' (3025); Gen. Pl. *siekt* (905) 'seckts' vs. *projekt* 'project' (51215); Gen.Pl. *jácht* (336) 'yachts' or *šácht* (317) 'pits' vs. Nom.Sg. *necht* (206) 'nail'.

There are also purely morphophonotactic word-final obstruent clusters, which originate only in Gen.Pl.: *ct* in *pôct* (325) 'of honours', *-pch* in *zápch* (90) 'obstructings' and *-rst* in *vierst* (211) 'wersts'.

In German, there are many word-final morphophonotactic obstruent clusters due to verb inflection, especially 2nd and 3rd Sg. or 2nd. Pl., in combination of *ch* /x/ with *st* or *t*: e.g. only morphophonotactic *ch-st* from *machen* 'make' in 2nd Sg. *mach-st*, whereas *ch-t* in 3rd Sg and 2nd Pl. *mach-t* has a phonotactic correspondence in *Macht* 'power'. In contrast, Slovak has such morphophonotactic word-final obstruent clusters only in Gen.Pl., as for example in Gen.Pl. *jácht* 'of yachts'. Thus, word-finally, German is richer in obstruent clusters than Slovak. Such obstruent clusters occur morphophonotactically in the German verbs: e.g.

flitz-t ‘flits’, *herz-t* ‘caresses’, *lach-t* ‘laughs’, *schaff-t* ‘creates’, *park-t* ‘parks’, *klapp-t* ‘flaps’, *lieb-t* ‘loves’, *verschlamp-t* ‘loses by neglect’, *lass-t* ‘lets’, *schirr-st* ‘harness’ (2nd Sg.), *schiel-st* ‘squint’, *schien-st* ‘put in splints’, *salb-st* ‘anoint’, *kerb-st* ‘notch’, *hack-t* ‘hacks’, 2nd Sg. *hack-st*, *krank-t* ‘is sick’.

German phonotactic correspondences of these word-final obstruent clusters occur in *Schaft* ‘shaft’, *Adept* ‘adept’, *First* ‘ridge’, *acht* ‘8’, *Arzt* ‘physician’, *prompt* ‘prompt’, *Wulst* ‘bulge’, *Wanst* ‘belly’, *Markt* ‘market’, *Papst* ‘pope’, *Herbst* ‘autumn’, *Takt* ‘tact’, *Text* ‘text’, *Punkt* ‘point’. Thus phonotactic word-final obstruent clusters occur in both languages, but they have a much higher type and token frequency in German than in Slovak, the opposite of what holds for word-initial obstruent clusters.

Only morphonotactic are the German word-final obstruent clusters in *salb-t* ‘anoints’, 2nd Sg. *salb-st*, *rupf-t* ‘plucks’, 2nd Sg. *rupf-st*, *welk-t* ‘withers’, 2nd Sg. *welk-st*, only 2nd Sg. *lach-st* ‘laugh’, *krank-st* ‘(you) are sick’.

Probably, segmentally identical phonotactic and morphonotactic clusters have different vowel durations (cf. Plag 2013), but it is yet unclear whether these differences lie above the threshold of perceptibility.¹ In any event, Plag is right in objecting to linguistic models which crucially contain a flow-chart from one submodule to another one in a way which presupposes bracket erasure (criticised also in Brown & Hippisley 2012: 273). Our model of morphonotactics (Dressler & Dziubalska-Kołaczyk 2006, Dressler *et al.* 2010, Korecky-Kröll *et al.* 2014) does not presuppose such bracket erasure. This also fits the above mentioned Slovak word-medial patterns: assuming that in a flow-chart inflectional morphology follows derivational morphology, the derivational boundary in *po-tok* must not be erased in order to prevent vowel deletion in Gen.Sg. *po-tok-a/u*, in contrast to the oblique cases of *líst-ok* and *otec*.

5. Phonotactic markedness

For German and other languages, there exist phonotactic analyses which assume any third consonant in a tautosyllabic consonant cluster as extrasyllabic or extrametrical (see Wiese 1988, 2000), which makes such a cluster a marked one. Additional consonantal morphemes increase markedness.

For our purpose, one of the most extensive and sophisticated markedness analyses can be found in Dziubalska-Kołaczyk’s (2002,

2009, Dziubalska-Kołaczyk & Zielinska 2010) Beats-and-Binding model, which goes beyond numerical complexity and purely sonority based models. This phonotactic theory measures the degree of binding of the consonants to each other and to beats (vowels and syllabic sonorants) by means of Net Auditory Distance (NAD). Being calculated as the distance which occurs between two segments in terms of manner and place of articulation, NAD evaluates word-initial and word-final consonant clusters as the more preferred the greater the relative NAD is between the beat and the adjacent more peripheral consonant in comparison to the NAD obtained between the peripheral consonants themselves.

A basic hypothesis in this model assumes that morphology renders phonotactics more marked. Thus morphonotactic obstruent clusters should be more marked than comparable phonotactic clusters (in the narrow sense). And indeed, if we compare the preferred cluster-rich positions of Slovak and German, then in Slovak word-initial positions the five most frequent morphonotactic biconsonantic obstruent clusters *s-č-*, *z-v-*, *v-ž-*, *s-ch-*, *z-ž-* are more marked than the five most frequent purely phonotactic obstruent clusters *sv-*, *št-*, *dv-*, *tv-*, *šk-*.

As to German word-final consonant clusters, the purely morphonotactic triconsonantal clusters *-f-st*, *-pf-st*, *-x-st* are among the most marked word-final clusters (cf. Korecky-Kröll *et al.* 2014). However more analyses of other clusters are needed.

6. On the acquisition of German (mor)phonotactic obstruent clusters

This psycholinguistic section is devoted to the relationship between phonotactic and morphonotactic consonant clusters in first language acquisition, because phonotactic markedness has to do with how patterns are acquired by young children. Freiburger (2007) investigated morphonotactic and phonotactic clusters in the longitudinal spontaneous speech data of a Viennese boy (aged 1;6 to 3;0) and, in contrast to Zydorowicz (2009) on Polish and English and to Kamandulyte (2006) on Lithuanian, did not find a clear advantage for the acquisition of morphonotactic as opposed to phonotactic clusters, but also no disadvantage. This can be interpreted as showing that in morphology-rich languages, such as the Slavic languages, the presence of a grammatical morpheme within a word-final, word-medial or word-initial consonant cluster helps phonological acquisition (see the notion of bootstrapping, cf. Weissenborn & Höhle 2001), because children are more stimulated to focus on the acquisition of a richer than of a poorer morphology (Xanthos *et al.* 2011).

Unfortunately no acquisition data are so far available for Slovak.

Here, we briefly present the results of an analysis of spontaneous speech data from 16 typically developing Viennese children (aged 2;11 to 3;6) and their main caretakers. There were two 30-minute-recordings per child, thus 16 hours of speech in total. Our analysis is limited to *-Cst* clusters, which are the most frequent triconsonantal clusters in German.

In contrast to the Austrian German electronic corpora and to the Slovak National corpus of written adult speech, we find a much higher percentage of morphonotactic than of phonotactic consonant clusters in the children and caretakers speech: 85 % of the children's and 90 % of the parents' word tokens with word-final *-Cst* clusters contain morphotactic boundaries. A frequency ranking of word tokens containing *-Cst* clusters shows that the top three clusters in child speech (CS) and in child-directed speech (CDS) are all morphonotactic and the same in children and parents, namely final morphonotactic *-n-st*, *-l-st* and *-k-st*. These are only 2nd sg. verbal forms, such as *du kann-st* 'you can', *will-st* 'want' or *mag-st* 'like' which are extremely important and frequent words in parent-child interaction. Apart from the top three clusters, CDS and CS are slightly different. In CDS the final morphonotactic cluster *-x-st*, as in *du brauch-st* 'you need', is more frequent, whereas *-f-st*, as in *du darf-st* 'you may' is more frequent in CS. In the fifth place we find word-final morphonotactic *-m-st*, as in *du komm-st* 'you come'. In CDS and CS the top phonotactic cluster is *-nst*, as in *sonst* 'otherwise'. Phonotactic *nst* is also the most frequent medial cluster of our consonant clusters, but in comparison to the most frequent morphonotactic clusters it still is very rare.

The interaction between parents and their young children represents a very specific genre of oral speech, which abounds in second singular verb forms. This shows that the distribution of morphonotactic vs. phonotactic consonant clusters may be very genre-specific. But clearly for the acquisition and establishment of markedness relations oral CDS is the most important speech genre.

7. Conclusion and outlook

Phonotactic asymmetries between word-initial, word-final and word-medial positions are well-known. This starts with how the universal preference for CV structures (Dziubalska-Kołodziej 2002, 2009) is realized in the three positions and depending on whether a word is monosyllabic, disyllabic or polysyllabic. What is interesting for the

contrastive analysis of Slovak and German obstruent clusters is their polar difference in the asymmetry between word-initial and word-final position: whereas Slovak is much richer in type and token frequencies of word-initial obstruent clusters, German is much richer in type and token frequencies of word-final obstruent clusters. Type frequency asymmetries proved to be radicalised in token frequency differences, which means that the dominant patterns are more profitable. These asymmetries hold both for phonotactic clusters in the narrow sense and for the morphonotactic ones, even for purely morphonotactic obstruent clusters. This basic difference between the two languages appear to go back to prehistoric or early historic major vowel deletions in German word-final positions as opposed to late Protoslavavic deletions of ultra-short jer and jor in preceding word positions.

In word-medial positions both languages have a rich array of phonotactic obstruent clusters (not analysed in this contribution) and of morphonotactic clusters originating in derivational morphology. There is however a strong contrast in compounding which creates morphonotactic clusters in German, but avoids them in Slovak.

Slovak is a more consonantal language than German (27 vs. 21), and allows the phonemes /r, l/ as stressed syllabic nuclei. This general profile is presumably also responsible for the greater amount of consonant clusters, which include also more marked clusters than German. Morphology favours this consonantal abundance in Slovak due to frequent vowel deletion, but disfavours it word-internally in compounding via vocalic interfixes (in contrast to German consonantal interfixes). But this is less relevant, because German has much more productivity in compounding than Slovak.

What makes a difference between phonotactic and morphonotactic obstruent clusters is the greater markedness of morphonotactic than of phonotactic clusters. Nevertheless the interaction of morphology and phonology has been found to favour morphonotactic clusters in processing and in the acquisition by typically developing children (as opposed to children with Specific Language Impairment), at least in the morphology-rich languages Polish and Lithuanian. Evidence for German acquisition is still unclear, and there is not yet any study for the acquisition of Slovak.

Nevertheless also the evidence from our huge electronic corpora of adult speech must be considered with caution. It is not entirely clear whether these data are representative for the languages analysed. All three electronic corpora are limited to written data produced by adults for adult readers (adult-directed speech). Moreover, the Austrian Media Corpus is restricted to journalistic prose, whereas the much smaller

Austrian Academy Corpus includes also literary, political, and scientific texts (including scientific popularisation). In contrast, the Austrian acquisition data at our disposal are oral data, and for adult productions limited to child-directed speech. For all these reasons, we plan to extend our analyses to adult-directed adult oral speech.

Addresses of the Authors

Wolfgang U. Dressler, Universität Wien, Institut für Sprachwissenschaft, Sensengasse 3a, Stiege, 3. Stock, A-1090 Wien & Institut für Corpuslinguistik und Texttechnologie der Österreichischen Akademie der Wissenschaften, Sonnenfelsgasse 19, 1010 Wien, Austria <wolfgang.dressler@univie.ac.at>

Miroslava Hliničanová, Katedra germanistiky, Filozofická Fakulta UMB, Tajovského 40, 974 01 Banská Bystrica, Slovakia <miroslava.hlinicanova@umb.sk>

Matej Ďurčo, Austrian Centre for Digital Humanities, Österreichische Akademie der Wissenschaften, Sonnenfelsgasse 19, A-1010 Wien, Austria <matej.durco@assoc.oeaw.ac.at>

Karlheinz Mörth, Austrian Centre for Digital Humanities, Österreichische Akademie der Wissenschaften, Sonnenfelsgasse 19, A-1010 Wien, Austria <Karlheinz.Moerth@oeaw.ac.at>

Katharina Korecky-Kröll, Institut für Sprachwissenschaft Universität Wien, Sensengasse 3a, A-1090 Wien, Austria <katharina.korecky-kroell@univie.ac.at>

Notes

¹ In the tradition of generative Lexical Phonology, Rubach (1993) assumes also for the synchronic analysis of today's Slovak the underlying phonemes *jer* and *jor*, which never occur on the surface, an abstract analysis which we reject (as already implicit from Postal's 1968 naturalness principle).

² In an ongoing project, we are currently testing the differences of vowel duration before phonotactic vs. morphonotactic consonant clusters, trying to fill in this research lacuna.

Bibliographical References

- Bauer Laurie 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
Berg Thomas 2014. On the Relationship between Type and Token Frequency. *Journal of Quantitative Linguistics* 21(3). 199-222.
Brown Dunstan & Andrew Hippisley 2012. *Network Morphology*. Cambridge: Cambridge University Press.

- Bybee Joan 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Dressler Wolfgang U. 1985. *Morphonology*. Ann Arbor: Karoma Publishers.
- Dressler Wolfgang U. & Katarzyna Dziubalska-Kořaczyk 2006. Proposing morphonotactics. *Italian Journal of Linguistics* 18. 249-266.
- Dressler Wolfgang U., Dziubalska-Kořaczyk Katarzyna & Lina Pestal 2010. Change and Variation in Morphonotactics. *Folia Linguistica Historica* 31. 51-67.
- Du Lifang & Xuezhong Zhang 2010. A Survey of the Measurements of Morphological Productivity. *English Language Teaching* 3. 60-63.
- Dvonč Ladislav and collective 2013. *Pravidlá slovenského pravopisu*. Bratislava: Veda.
- Dziubalska-Kořaczyk Katarzyna 2002. *Beats-and-Binding Phonology*. Frankfurt am Main: Lang.
- Dziubalska-Kořaczyk Katarzyna 2009. NP extension: B&B phonotactics. *PSiCL* 45. 55-71.
- Dziubalska-Kořaczyk Katarzyna & Jarosław Weckwerth 2002. *Future Challenges for Natural Linguistics*. Munich: Lincom.
- Dziubalska-Kořaczyk Katarzyna & Daria Zielińska 2010. Predicting phonotactic difficulty in second language acquisition. In Rauber Andreia S., Watkins Michael A., Silveira Rosane & Rosana D. Koerich (eds.), *The acquisition of second language speech: Studies in honor of Professor Barbara O. Baptista*. Florianópolis: Editora Insular. 281-304.
- Freiberger Eva Maria. 2007. Morphonotaktik im Erstspracherwerb des Deutschen. *Wiener Linguistische Gazette* 74. 1-23.
- Kilani-Schoch Marianne & Wolfgang U. Dressler 2005. *Morphologie naturelle et flexion du verbe français*. Tübingen: Narr.
- Korecky-Kröll Katharina, Dressler Wolfgang U., Freiberger Eva Maria, Reinisch Eva, Mörth Karlheinz & Gary Libben 2014. Morphonotactic and phonotactic processing in German-speaking adults. *Language Sciences* 46. 48-58.
- Krajčovič Rudolf 1971. *Náčrt dejín slovenského jazyka*. Bratislava: Slovenské pedagogické nakladateľstvo.
- Plag Ingo 2013. Against homophony: the acoustics of English {s} morphemes. Paper presented at the 3rd NetWordS workshop, Dubrovnik, September 2013.
- Postal Paul 1968. *Aspects of Phonological Theory*. New York: Harper and Row.
- Rubach Jerzy 1993. *The Lexical Phonology of Slovak*. Oxford: Clarendon Press.
- Růžička Jiří and collective 1966. *Morfológia slovenského jazyka*. Bratislava: Vydavateľstvo Slovenskej akadémie vied.
- Sabol Ján 1989. *Syntetická fonologická teória*. Bratislava: Veda.
- Sokolová Miloslava, Moško Gustáv, Šimon František & Vladimír Benko 1999. *Morfematický slovník slovenčiny*. Prešov: Náuka.
- Tomasello Michael 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge: Harvard University Press.
- Weissenborn Jürgen & Barbara Höhle 2011. *Approaches to Bootstrapping*:

Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition. Amsterdam: Benjamins.

Wiese Richard 1988. *Silbische und Lexikalische Phonologie.* Tübingen: Niemeyer.

Wiese Richard 2000. *The Phonology of German.* Oxford: Oxford University Press.

Wurzel Wolfgang U. & Manfred Schentke 1989. *Inflectional Morphology and Naturalness.* Dordrecht: Kluwer.

Xanthos Aris, Laaha Sabine, Gillis Steven, Stephany Ursula, Aksu-Koç Ayhan, Christofidou Anastasia, Gagarina Natalia, Hrzica Gordana, Ketrez F. Nihan, Kilani-Schoch Marianne, Korecky-Kröll Katharina, Kovačević Melita, Laalo Klaus, Palmović Marijan, Pfeiler Barbara, Voeikova Maria D. & Wolfgang U. Dressler 2011. On the role of morphological richness in the early development of noun and verb inflection. *First Language* 31. 461-479.

Sonority and early words: the Sonority Syllable Model applied to an acquisitional project with Danish children

Laila Kjærbæk, Ditte Boeg Thomsen, Claus Lambertsen & Hans Basbøll

Syllables play an important role in children's early language acquisition, and children appear to rely on clear syllabic structures as a key to word acquisition (Vihman 1996; Oller 2000). However, not all languages present children with equally clear cues to syllabic structure, and since the specific L1 is known to influence the acquisition rate, it is worth investigating how sonority characteristics in children's linguistic input influence their lexical acquisition. Danish is characterized by having a monotonous sonority envelope compared to other Scandinavian languages. Danish language acquisition therefore presents us with the opportunity to examine how children respond to the task of word learning when the input language offers less clear cues to syllabic structure than usually seen. To investigate the sound structure in Danish children's lexical development, we need a model of syllable structure; and as the theoretical basis for our analyses related to sonority we present Basbøll's Sonority Syllable Model for phonotactics, which is based upon a non-circular version of a sonority hierarchy. We investigate spontaneous child language output in a longitudinal corpus with two children aged 9-29 months. For the two children, the phonetic structure of the first ten words to occur is compared with that of the last ten words to occur before 30 months of age, and with that of ten words in between. Measures related to the sonority envelope, viz. sonority types and in particular sonority rises, are applied to this data.

1. Syllables in children's language acquisition

In children's early language acquisition, the structure of canonical syllables with transitions between consonants and vowels appears to be of major importance to children's parsing of their input and to their own productions. The onset of canonical babbling with well-formed and temporally regular syllables is often regarded as "the most striking production milestone in the first year" (Vihman 1996: 109), and the occurrence, amount and types of CV syllables in babbling is predictive of many later developments in language acquisition (Kent 1992, Stoel-Gammon 1998). Further, children have been found to rely on the specific canonical syllables 'practiced' during babbling in their earliest word productions (Vihman 1992). Oller (2000) suggests that syllables with their relatively consistent timing are fundamental to children's parsing of the input stream because they

impose rhythmical structure on speech and “provide a frame of focus for perception, allowing the listener to perform analysis on limited chunks of information” (Oller 2000: 91).

We assume that children develop a network of sound-based associations, cross-indexing the words they know regarding for example syllable-count and matching initial or final syllables (Vihman 1981: 247). The stressed syllable seems to be of significant importance – yet we might expect the syllable to be more important in stress-timed languages like English than in more syllable-timed (‘weakly stressed’) languages like French (Vihman 1981: 254). Danish is expected to be like English in this respect.

2. The Sonority Syllable Model

2.1. Vocoid: the prototypical peak of a syllable

All languages of the world can be described using syllables as structures.¹ All languages have vowel segments which form the peak of a syllable. Some, but not all, languages also have non-vowels (consonants) that can form the peak of a syllable.² All languages of the world have non-vowels that do not form the peak of a syllable. Some, but not all, languages also have vowels that do not form the peak of a syllable.³ Thus the prototypical peak of a syllable is phonetically a vowel segment, a “vocoid”, whereas the prototypical non-peak of a syllable is phonetically a consonant, a “contoid” (Pike’s terminology 1943).

Basbøll (since Basbøll 1973) has used an equation as the definition of vocoid:

$$[\text{vocoid}] =_{\text{DEF}} [\text{sonorant}, -\text{stop}, -\text{lateral}]$$

2.2. Universal logic of segment types: independent of time and segment order

All vocoids are, necessarily, sonorant.⁴ But some sonorants are not vocoids, viz. prototypical (sonorant) laterals and nasal consonants (which are [sonorant, stop]). Furthermore, all sonorants are, necessarily, voiced.⁵ On the other hand, there are non-sonorant sounds (obstruents) that are voiced. Also voiceless segments occur, of course. This universal logic of segment types, built exclusively on existing vs. excluded segment types, can be formulated as an implication and be depicted by a set of (concentric) Euler’s circles (Figure 1):

$$[\text{vocoid}] \text{ implies } [\text{sonorant}] \text{ implies } [\text{voiced}]$$

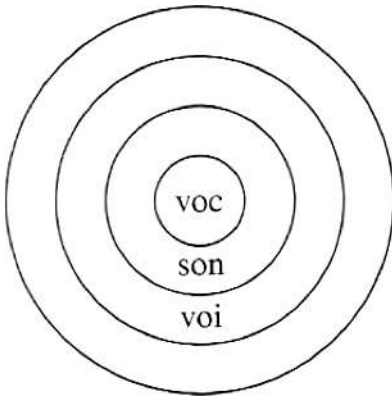


Figure 1. A set of (concentric) Euler's circles depicting the logical relation between vocoids, sonorants and voiced segments independently of any time- or order aspect (from Basbøll 2005: 194).

2.3. Introducing time: a syllable model emerges

When the dimension of time is introduced into the set of concentric circles, a set of relevant order classes of phonotactics follows, predicting the order (up to the peak): voiceless segments, voiced obstruents, sonorant contoids, vocoids; and the mirror-image order in the final part of the syllable. Thus the introduction of the time dimension turns the model into a Sonority Syllable Model (Figure 2):

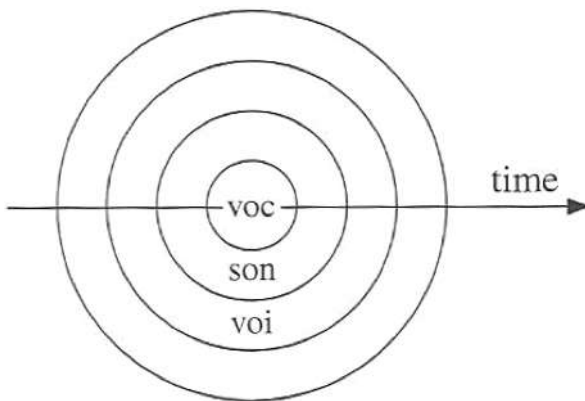


Figure 2. The Sonority Syllable Model. Identical to the set of Euler's circles depicted in Figure 1, but with the addition of the dimension of time which turns the figure into a syllable model (from Basbøll 2005: 184).

2.4. Margins of the Sonority Syllable Model: [spread glottis]

One further phonologically relevant distinctive feature satisfies the criteria for entering into the implication chain underlying the model: [spread glottis], which we understand as meaning 'widely spread glottis'. There are segments that are voiceless but do not have

(widely) spread glottis.⁶ This means that the maximal version of the Sonority Syllable Model can be based upon the following implication chain:

[vocoid] implies [sonorant] implies [voiced] implies [–spread glottis]

One consequence of this implication chain in the Sonority Syllable Model is that its most marginal segments have the feature [spread glottis]. This makes very good sense phonetically, in our view: the isolated monosyllable, which is the prototypical isolated syllable, starts and ends in rest position which has, obviously, widely open passage through the glottis (for breathing).⁷

2.5. Conclusion on the Sonority Syllable Model

The Sonority Syllable Model predicts that sequences specified in terms of sonority types exhibit ‘mirror image’ structure initially and finally. This is generally true. But it does not follow from the Sonority Syllable Model that sequences of equal sonority are ‘mirror image’-like. On the contrary, they strongly tend to have the same order initially and finally, following separate principles, and the distinction between one and a sequence of two identical segments must be found outside sonority, viz. in prosody and in substitutability.⁸

3. Some phonological differences within Scandinavian: Danish and the sonority envelope

As captured by the Sonority Syllable Model, syllable structure in the languages of the world is fundamentally characterized by rises and falls in sonority. Since the syllable appears to be an important entity in early child language acquisition, we expect clear differences in sonority to facilitate children’s parsing of the input stream and their lexical acquisition. However, the languages of the world of course differ widely as to phonotactics, i.e. the concrete syllable structures allowed by the specific languages and the Sonority Syllable Model’s circles activated, and also as to which types of reductions characterize spontaneous speech. So, even if children can be generally expected to be highly sensitive to rises and falls in sonority in the input stream during early lexical acquisition, the languages of the world present them with very different cues to syllable structure. Danish is characterized by an unusually monotonous sonority envelope and may therefore present L1 learners with special challenges.

This becomes particularly clear when we compare Danish and its close Scandinavian relatives, in particular its closest relative Swedish (which is also East Nordic).

In Danish, coda lenition, light stressed syllables, schwa elision, /r/ elision, semi-vowel elision and length vacillation are all features causing indistinct sound structure, whereas in most forms of Swedish, word accents, final lengthening and compulsory sentence accent are all strong prosodic cues making sound structure more distinct (Grønnum 2003, 2005), see Table 1. Therefore Swedish is most likely easier for children to grasp.

Table 1. Sound structure in Danish and Swedish (from Basbøll 2012: 24, based upon Grønnum's work, see Grønnum 2003, 2005, 2008).

	DANISH	SWEDISH
Coda lenition	Yes	No
Light stressed syllables	Yes	No
Schwa elision	Yes	No
/r/ elision	Yes	No
Semi-vowel elision	Yes	No
Length vacillation	Yes	No
Tonal word accents	No	Yes
Final lengthening	No	Yes
Compulsory sentence accent	No	Yes
Signal for utterance function	Weak and global	Strong and local

Corresponding to Swedish (non-sibilant) obstruents, Danish has (non-lateral) approximants and glides, i.e. segments which belong to the same sonority class as vowels, viz. vocoids. There is a combined effect of weakening and schwa-reduction, for example Swedish *gata* [²ga:ta] and *koka* [²kʰu:ka] vs. Danish *gade* [ˈgæ(:)ð(:)] and *koge* [ˈkʰo(:)w(:)].⁹

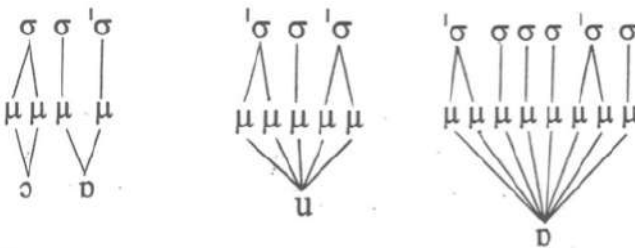


Figure 3. The syllable structure in Danish *koge* o(ver) ‘boil over’, (l)uge *u*(denfor) ‘weed outside’, (h)årdere at åre(lade) ‘harder to bleed’ (Rischel 2003: 279).

In Figure 3 Rischel (2003) demonstrates that the structure of the Danish syllable (which he calls ‘a national heritage’) leads to a syllabic-moraic structure where a single vowel can realize a sequence of morae and syllables.¹⁰

Table 2 exhibits the past tense forms of a fully regular and fully productive weak verb, viz. Danish *love* ‘promise’ and its Scandinavian cognates in Norwegian (bokmål¹¹), Swedish and Icelandic. The most significant observation comes from the column ‘Sonority rises from the stem-final consonant’. In Danish there is no indication whatsoever in the sonority envelope that a syllable (containing the past time morpheme) has been added to the stem, cf. Figure 3. The phonology, in particular sonority and syllable structure, thus makes the morphological structure of such forms in Danish far more opaque than is the case in our close Scandinavian relatives.

Table 2. Weak past tense forms in Scandinavian (based upon the principles of Bleses et al. 2011: 16). Example: past tense of (Danish) *love* ‘promise’ and its Scandinavian cognates (from Basbøll 2012: 26).

	SPOKEN FORM	NUMBER OF SYLLABLES IN SUFFIX (0,1,2)	NUMBER OF VOWELS IN SUFFIX (0,1,2)	NUMBER OF SONORITY RISES FROM THE STEM-FINAL CONSONANT (0,1,2)	WORD ACCENT CUE FOR SUFFIX (NON- STØD/TONEME 2) (0,1)
Danish					
-ede	[ˈlɔːʊð]	2	0	0	(1)
	[ˈlɔːʊð]	1	0	0	(1)
Norwegian					
-et	[²lɔːvæt]	1	1	1	1
-a	[²lɔːva]	1	1	1	1
Swedish					
-ade	[²lɔːvadə]	2	2	2	1
-a	[²lɔːva]	1	1	1	1
Icelandic					
-aði	[ˈlɔːvaði]	2	2	2	0
-aðir	[ˈlɔːvaðɪr]	2	2	2	0
-uðum	[ˈlɔːvʏðʏm]	2	2	2	0
-uðuð	[ˈlɔːvʏðʏð]	2	2	2	0
-uðu	[ˈlɔːvʏðʏ]	2	2	2	0

4. Hypotheses

In perception canonical syllables with clear transitions between consonants and vowels are privileged units for children to analyze (Oller 2000). For their own production of canonical syllables in pre-linguistic babbling children depend on their ambient language (Oller & Eilers 1988), and these babbling syllables later serve them as bases for their own first words (Vihman 1996). Danish presents children with fewer cues to syllable structure than many other languages do, and this makes early Danish lexical development particularly interesting. The specific sonority characteristics of the input language could be expected to affect children's selection of target words for production, their actual pronunciation of the selected target words and/or acquisition rate. In this study, we focus on the first two issues (with the exclusion, therefore, of acquisition rate) and investigate Danish children's early words across three age stages in terms of segments and syllables per word, sonority types and number of sonority rises, examining both target word selection and actual pronunciation. For all these measures (segments, syllables, sonority types and sonority rises per word), we expect to see a rise at each of our measure points, but we cannot predict whether such expansion will show up in children's selection of words, in their pronunciation of words (e.g. less reduction with age) or in both. We do not consider these measures to be language specific, i.e. we expect a developmental increase in these measures in other languages as well. But we also expect the exact developmental evolution with respect to these measures to be dependent on language type. Further, we will seek to assess whether Danish children prefer words with clear syllable structure to words with vague syllable structure as targets for pronunciation, and whether such potential preference/avoidance patterns are equally active across development within the time window investigated.

5. Method and empirical data

In order to test our hypothesis we have analyzed longitudinal data from the Odense Twin Corpus, a corpus of spontaneous child language input and output. We used data from a dizygotic twin pair, namely the girls Ingrid and Sara in the age of 9-29 months. The children and their family were recorded for about an hour once every month. All recordings are transcribed in CHILDES (MacWhinney 2000 a,b) in Danish standard orthography and coded in OLAM¹², where all words were provided with morphological information as well as their standard pronunciation (cf. Kjærbæk 2013).

We know from earlier studies that children's pronunciations vary more than adults' from one time to the next, and, in their first few months of talking, they may produce multiple versions of the same word (Maekawa & Storkel 2006; Sosa & Stoel-Gammon 2006). In this study we focus on a minor set of words, which we will then follow throughout the time window. This gives us the opportunity to study the development of the specific word during time as well as the child's different pronunciations of the same word in the same recording.

For each child we have registered: 1) the first ten words in the corpus (A-words); 2) the last ten words in the time window (9-29 months) (C-words); 3) ten words exactly in the middle (B-words). All occurrences of the A-, B- and C-words are transcribed phonetically in child speech of the respective child as well as in her parental input (Basbøll et al. 2012a,b), see Tables 3-5.

To decide whether sound strings should be counted as babbling or first words (i.e. A-words), we followed the explicit word identification procedure in Vihman & McCune (1994), evaluating each possible word candidate on the following criteria: Context (determinative context, caretaker identification, multiple use, multiple episodes), Vocalization shape (complex match, exact match, prosodic match), Relation to other vocalizations (imitated tokens, invariant shape, no inappropriate uses).

Table 3. Ingrid and Sara's A-words including standard adult pronunciation and the child's pronunciation.

INGRID				SARA			
WORD	GLOSS	STANDARD	CHILD	WORD	GLOSS	STANDARD	CHILD
<i>mmm</i>	(tastes-good sound)	['m:m:]	['m:m:]	<i>mmm</i>	(tastes-good sound)	['m:m:]	['m:m:]
<i>nam</i>	(tastes-good sound)	['nam]	['am:]/ ['nam]	<i>nej</i>	'no'	['najʔ]	['ŋaʔ]
<i>ja</i>	'yes'	['ja]	['ja]	<i>muh</i>	'moo'	['mu:]	['u:u:]
<i>op</i>	'up'	['ʌb]	['aʔ]	<i>vov</i>	'woof/bow-wow'	['vʌw]	['ʌw]
<i>nej</i>	'no'	['najʔ]	['na:jʔ]	<i>mam</i>	(child form for 'food')	['mam]	['ma]
<i>det</i>	'this/that'	['de]	['de]	<i>op</i>	'up'	['ʌb]	['ʌb]
<i>der</i>	'there'	['dæɐ̯]	['deɐ̯h]	<i>ah</i>	(tastes-good sound)	['æ:]/['a:]	['a:h]/ ['a:a:]
<i>mælk</i>	'milk'	['mɛlʔg̊]	['mɛ:]	<i>uhm</i>	(tastes-good sound)	['ɔm]	['ɔm:]
<i>se</i>	'look'	['se:ʔ]	['se:e:]	<i>nam</i>	(tastes-good sound)	['nam]	['nam]
<i>mam</i>	(child form for 'food')	['mam]	['ma]/ [ma:m]	<i>mælk</i>	'milk'	['mɛlʔg̊]	['m:ɛ]

6. Results

6.1. Segments per target word

We now present some quantitative results from comparing Ingrid and Sara's A-, B- and C-words. First, we have counted the number of segments per target word (A-, B- and C-words separately as well as collapsed) in Ingrid and Sara's target words. The results are shown in Table 6. The last column summarizes the measure for all target words taken together.

Table 6. Number of segments¹³ per target word (10 A-words, 10 B-words and 10 C-words). Standard pronunciation (I) indicates the standard pronunciation of Ingrid's target words and similarly for Sara: (S).

	A-WORDS	B-WORDS	C-WORDS	SEGMENTS PER TARGET WORD (AVERAGE A+B+C)
Ingrid (I)	2.5	3.7	3.9	3.4
Standard pronunciation (I)	2.6	4.2	5.5	4.1
Sara (S)	2.2	3.3	5.3	3.6
Standard pronunciation (S)	2.5	4.1	6.3	4.3
Children (I+S)	2.4	3.5	4.6	3.5
Standard pronunciation (I+S)	2.6	4.2	5.9	4.2

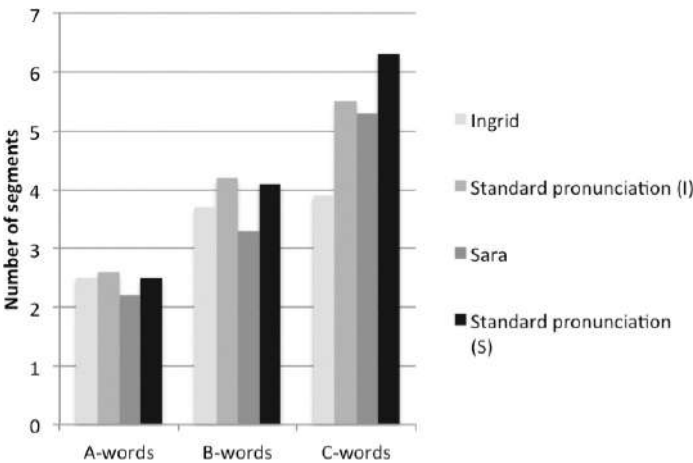


Figure 4. Number of segments per target word (10 A-words, 10 B-words and 10 C-words).

The results of Table 6 are depicted in Figure 4.

The two twin sisters exhibit a different developmental pattern with respect to the measure ‘segments per target word’: for both sisters, the A-words are simpler than the B- and C-words; but for Ingrid, there is only a small difference between B- and C-words, whereas for Sara, C-words are clearly longer than B-words. The same difference is seen in the standard pronunciation of the target words. Furthermore, we see that Ingrid and Sara’s pronunciation of the B- and C-words have fewer segments, i.e. they are shorter, than in the standard pronunciation. We also see a phonetic reduction in child speech when we compare it with the standard pronunciation.

6.2. Syllables per target word

Another possibility, still rather simple but perhaps more interesting in the present perspective, is to count length in terms of syllables rather than segments. The results of this analysis are shown in Table 7.

Table 7. Number of syllables¹⁴ per target word (10 A-words, 10 B-words and 10 C-words). Standard pronunciation (I) indicates the standard pronunciation of Ingrid’s target words and similarly for Sara: (S).

	A-WORDS	B-WORDS	C-WORDS	SYLLABLES PER TARGET WORD (AVERAGE A+B+C)
Ingrid (I)	1.2	1.6	1.7	1.5
Standard pronunciation (I)	1.0	1.6	2.2	1.6
Sara (S)	1.3	1.5	2.1	1.6
Standard pronunciation (S)	1.0	1.6	2.3	1.6
Children (I+S)	1.3	1.6	1.9	1.6
Standard pronunciation (I+S)	1.0	1.6	2.3	1.6

The results of Table 7 are depicted in Figure 5.

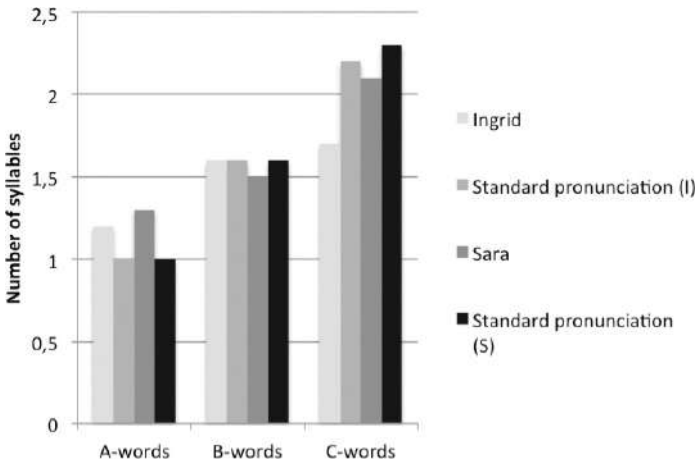


Figure 5. Number of syllables per target word (10 A-words, 10 B-words and 10 C-words).

The two twin sisters exhibit a different developmental pattern with respect to the measure ‘syllables per target word’, but the pattern is not the same as that found for segments (in Figure 4). For Ingrid, there is a rather small developmental evolution with respect to the number of syllables: B- and C-words are roughly at the same level (with a very slight increase). For Sara, by contrast, there is a big increase in length with respect to number of syllables per target word, from B-words to C-words. As for the pronunciation of the target words, we see a progressive increase for the two girls – Sara with a little higher amount of segments, syllables, sonority types and sonority rises than Ingrid.

The children mainly produce monosyllabic words in the beginning but as they reach the age of 29 months they generally produce disyllabic words. These results are interesting since studies have shown that disyllables dominate the early lexicon of children acquiring most of the other languages in which early word phonology has been investigated (Estonian, Finnish, French, Greek, Hebrew, Hindi, Italian, Japanese, Spanish, Welsh). The Germanic languages, though, generally seem to constitute exceptions, as monosyllables appear to be the most common early words in Dutch (e.g. Elbers & Ton 1985) and German (Leopold 1939; Elsen 1996) as well as English; for Swedish data show that mono- and disyllabic early word forms are in close balance (cf. Vihman & Croft 2013: 20).

A typical example of Ingrid's A-words is the monosyllabic word *ja* 'yes' ['ja] and a typical example of her C-words is the four syllable word *skraldespanden* 'the garbage can' which she reduces to a three syllable word ['gɑ:l, ɸan²-]. Generally, the children reduce more syllables in the C-words than in both the A- and B-words which is not surprising since there are more syllables to be reduced in C-words. According to the literature, children acquiring Germanic languages often omit weak or unstressed syllables in initial or final position in their early productions (e.g. Smith 1973). However, the fact that the A-words are very important for the communication since they always carry the whole meaning of the utterance (one word sentences), whereas the C-words are usually part of a longer utterance, and thus relatively less important, could also play a role.

6.3. *Different sonority types per target word*

The measures in sections 6.1 and 6.2 have given some indication of complexity in the trivial sense of length (measured in terms of number of segments and syllables). However, because of the particular character of Danish as having a monotonous sonority envelope and long vocoidal stretches, this measure needs to be complemented by a measure focusing upon the sonority envelope, i.e. sonority ups and downs. We propose two such measures here, viz. sonority types (in the present section) and sonority rises (in section 6.4; see also Table 2 above for preliminary information). Sonority types and sonority rises are both measures of complexity with respect to sonority, not just measures of length. Both are derived from the Sonority Syllable Model. It follows from this model that there are exactly five sonority types in the sense of major classes having different positions on a sonority hierarchy, neither more nor less: vocoids, sonorant contoids, voiced obstruents, voiceless obstruents with non-spread glottis, and, finally, obstruents with spread glottis.

We have counted the number of different sonority types per target word in Ingrid and Sara's target words. The results are shown in Table 8.

Table 8. Number of different sonority types¹⁵ per target word (10 A-words, 10 B-words and 10 C-words). Standard pronunciation (I) indicates the standard pronunciation of Ingrid’s target words and similarly for Sara: (S).

	A-WORDS	B-WORDS	C-WORDS	SONORITY TYPES PER TARGET WORD (AVERAGE A+B+C)
Ingrid (I)	1.9	2.3	2.5	2.2
Standard pronunciation (I)	1.9	2.9	3.2	2.7
Sara (S)	1.7	2.3	3.2	2.4
Standard pronunciation (S)	1.9	2.6	3.5	2.7
Children (I+ S)	1.8	2.3	2.9	2.3
Standard pronunciation (I+S)	2.2	2.8	3.4	2.7
Children (I+ S)	0.1	0.0	-0.35	-0.08
Standard pronunciation (I-S)	0.0	0.15	-0.15	0.0

Ingrid and Sara grew up in the same environment, and we therefore consider the study of their differences and similarities particularly interesting. In order to highlight similarities and dissimilarities, besides individual scores, we have also considered the average scores for their pronunciation (I + S), as well as the difference between their pronunciation and the corresponding standard pronunciation of the two girls (I – S); the largest difference is found for the C-words, where Sara represents the highest number of sonority types whereas Ingrid leads for the A- and B-words, but with a much smaller margin.

The results from Table 8 are depicted in Figure 6.

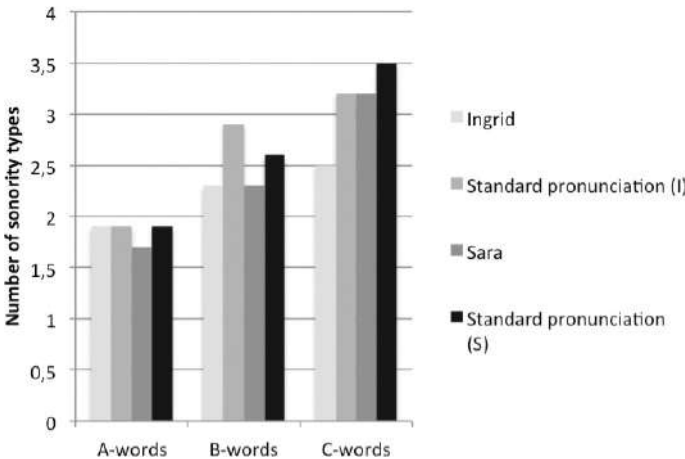


Figure 6. Number of different sonority types per target word (10 A-words, 10 B-words and 10 C-words).

Again we see that the two twin sisters exhibit a slightly different developmental pattern with respect to the measure 'different sonority types per target word': for both sisters, the A-words have less sonority types than the B-words which again have less sonority types than the C-words. This measure of sonority thus presents a stronger developmental evolution than the previous measures (number of segments, and of syllables, per target word). There is a small difference between the two twin sisters, however, inasmuch as the increase from B-words to C-words is visibly larger for Sara than for Ingrid. This is perhaps a reflection of the same pattern as in the previous section: at the last stage, Sara produces a few longer words than Ingrid, and this gives her the opportunity to use more different sonority types and (as we shall see in the next section) sonority rises. Since we only analyze ten words per child, we cannot expect this difference between the two sisters' target C-words to be representative of more general differences, nor can the difference be considered representative of general patterns in Danish children given the limited number of informants. We can be a little more confident in the difference found between age stages because both children exhibit the same development.

6.4. Sonority rises per target word

To obtain a more revealing measure for the diversity in the sonority envelope than the mere number of different sonority types per word (as in section 6.3), we have calculated the number of sonority rises per target word (cf. Table 2) for Ingrid and Sara. The results are shown in Table 9.

Table 9. Number of sonority rises¹⁶ per target word (10 A-words, 10 B-words and 10 C-words). Standard pronunciation (I) indicates the standard pronunciation of Ingrid’s target words and similarly for Sara: (S).

	A-WORDS	B-WORDS	C-WORDS	SONORITY RISES PER TARGET WORD (AVERAGE A+B+C)
Ingrid (I)	0.7	1.2	1.2	1.0
Standard pronunciation (I)	0.7	1.8	1.9	1.5
Sara (S)	0.3	1.2	1.8	1.1
Standard pronunciation (S)	0.6	1.3	2.4	1.4
Children (I+ S)	0.5	1.2	1.5	1.1
Standard pronunciation (I+S)	0.7	1.6	2.2	1.5
Children (I+ S)	0.2	0.0	−0.3	−0.03
Standard pronunciation (I−S)	0.05	0.25	−0.25	0.02

The results of Table 9 are depicted in Figure 7.

The two twin sisters exhibit a slightly different developmental pattern with respect to the measure ‘sonority rises per target word’ (Figure 7): for both sisters, the A-words have less sonority rises than the B-words which again have less sonority rises than the C-words. Examples for Sara are: A-word *mælk* ‘milk’ [‘m:ɛ] (standard pronun-

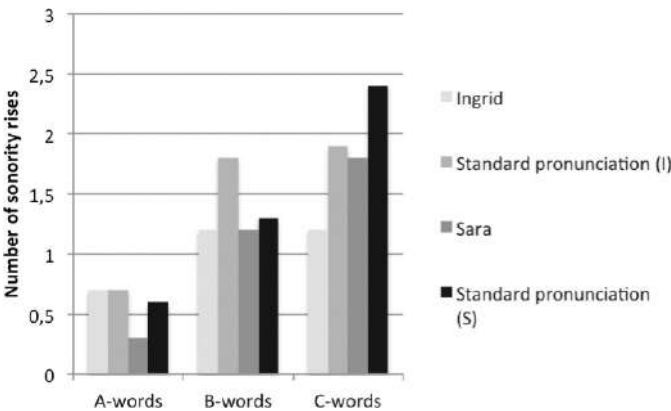


Figure 7. Number of sonority rises per target word (10 A-words, 10 B-words and 10 C-words).

ciation ['mɛl'g]); B-word *spiser* 'eat' ['gʁisɐ] (standard pronunciation ['sɕi:'sɐ]); and C-word *hundehvalp* 'puppy' ['hunə, val'ɸ] (standard pronunciation ['hunə, val'ɸ]) – all three words represent Sara's maximum in sonority rise among the target words at the three different stages. This measure of sonority thus presents a stronger developmental evolution than the first two measures (number of segments, and of syllables, per target word). There is a small difference between the two twin sisters, however, inasmuch as the developmental increase in this measure (sonority rises per target word) is larger (i.e. the 'curve' would be steeper) for Sara than for Ingrid. As for the standard pronunciation of the target words (Table 3-5), the difference in word length, number of segments, syllables, (different) sonority types and sonority rises is already present in regard to which words the children try to pronounce – even though we are not sure how distinctly they are pronounced by the parents.

7. Discussion and further perspectives

Earlier studies indicate that Danish children are late in their early language acquisition (e.g. lexicon and past tense inflection) compared to children acquiring other languages, but already at the age of 3 years Danish children seem to catch up with other children (Bleses et al. 2008). This could be an indication that the indistinct Danish sound structure is a challenge in the very early language acquisition, i.e. at the stage where the children are to acquire (identify, segment etc.) the speech sounds of the specific language. The Danish sound structure may be a challenge in relation to the acquisition of speech sounds of the Danish language, i.e. already at the early stage where the children are acquiring monosyllables. This is due, we claim, to two (still active) weakening processes which have conspired to make Danish phonetic structure opaque, viz. final obstruent weakening and schwa reduction. As a consequence, it is not always possible to locate the appropriate syllable boundaries within one smoothly developing sonority slope or even to determine the correct number of syllables. This furthermore contributes to the opacity of morpheme boundaries in Danish, much more so than in its closest relative, Swedish. Since clear syllable structure, and the rhythm of sonority rises and falls it imposes on speech, provide children with privileged units for analysis (Oller 2000), Danish L1 acquisition can give us an idea how lexical development proceeds in languages offering less rich cues to syllable structure.

As for children's selection of target words from their ambient language, what we have found is a constant rise from A-words to B-words to C-words for all four measures: The target A-words contain less segments, syllables, (different) sonority types and sonority rises than the target B-words, and these again contain less segments, syllables, (different) sonority types and sonority rises than the target C-words. Turning to children's actual pronunciations of target words, they generally contain fewer segments, syllables, sonority types and sonority rises than the standard pronunciation of the words, but the sonority measures for actual pronunciations follow the same expansion over the twenty months as seen in the selection of target words.

As for the question whether Danish children avoid target words with vague syllable structure in the ambient language, we can say that for the middle and final words in this study (B- and C-words), there is no general avoidance, the children for example selecting *noget* 'something' (standard pronunciation ['nɔ:əð]) and *over* 'over' (standard pronunciation ['ʌwʔɐ]). If we look at the two children's first words (A-words), however, these are characterized by the absence of words with unclear transition between syllables – because they are all monosyllables. This could be interpreted in different ways: either the children specifically avoid words with unclear syllable structure at this earliest stage of lexical acquisition, or the children in general simply prefer (stressed) monosyllables. In the latter case, according to our sonority measures proposed in this paper, there would not be a clear systematic difference between Danish and Swedish monosyllables. If we include the number of sonority falls (in addition to sonority rises), however, monosyllables in Swedish would be more complex than in Danish with respect to the sonority envelope (compare e.g. Danish *mad*, ending in a vocoid, and its Swedish cognate *mat*, ending in a plosive).

The Danish and Swedish data are not directly comparable (partly due to morphological differences), though, and future studies comparing early lexical acquisition in these two sister languages with more equivalent data sets would be useful for illuminating the influence of sonority characteristics of ambient language on children's first words. Similarly, it would be interesting to compare vocabulary development in Danish with the vocabulary development found in unrelated languages with equal sonority characteristics to assess whether the patterns of lexical selection we find in Danish acquisition are typical of languages with a relatively monotonous sonority envelope or not.

Our conclusions on the development of early Danish vocabulary should be read cautiously for at least two reasons. First of all, we analyzed data from only two children, a pair of twins. This means that we

cannot know how representative the patterns found are of early Danish vocabulary development in general, and the analysis should certainly be broadened in future studies. As a first step, we have begun analysis on a second pair of twins of the same age group, and the preliminary results suggest a replication of our findings here. Secondly, it is worth remembering that what we present as target pronunciations here are standard pronunciations as given in Danish dictionaries as norms for distinct pronunciation. As yet, we cannot be sure that children have actually encountered these distinct pronunciations in their input; in principle, what we present as children's reductions compared to input forms could in fact be full forms compared to what they actually hear. However, we have also transcribed all parents' actual phonetic realizations of all A-, B- and C-words; as a consequence, in new analyses we will be able to use those (the most distinct and/or the most frequent forms) as bases of comparison for child pronunciations. For the cases we have already checked, there is no doubt that children *do* reduce compared to input form, for example Ingrid's B-word ['*ɸis*] (*spise*, 'eat'), which corresponds to ['*s ɸi:s*] in her parental input.

It would be interesting to study further whether the indistinct Danish sound structure has more effect on the acquisition rate than on the children's selection of words. And to conclude, the measures of complexity with respect to sonority presented in this paper – based upon a general and non-circular model of sonority which defines exactly five sonority types – should be applied to other languages with different (genetic and) typological characteristics, as a potential contribution to a typology of sonority changes with respect to L1 acquisition.

Addresses of Authors

Laila Kjærbaek, Department of Language and Communication, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark
<kjarbak@sdu.dk>

Ditte Boeg Thomsen, Department of Scandinavian Studies and Linguistics, University of Copenhagen, Njalsgade 120, DK-2300 Copenhagen S, Denmark
<fjt414@hum.ku.dk>

Claus Lambertsen, Pestalozzistrasse 138, D-14612 Falkensee, Germany
<l.claus@lambertsen.de>

Hans Basbøll, Department of Language and Communication, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark
<hba@sdu.dk>

Acknowledgements

We would like to thank the Danish Council for Independent Research, Culture and Communication and the University of Southern Denmark for support. We are also grateful for the very useful comments that we received from two anonymous reviewers.

Notes

¹ Section 2 is based upon Basbøll (2012), section 3.

² Such as /l/ in the second syllable of English *little*, or in the only syllable of Czech *vlk* ‘wolf’.

³ Semi-vowels or glides, such as the first segment in the French monosyllable *oui* ‘yes’. It is a vowel, phonetically, despite its function as non-peak.

⁴ Sonorants are defined acoustically (following Ladefoged 1971: 58: “a comparatively large amount of acoustic energy within a clearly defined formant structure”; Ladefoged 1971: 93: “greater acoustic energy in the formants”).

⁵ This follows from the definition used here combined with the phonetic (articulatory and acoustic) fact that in order to get great acoustic energy in the formants, the vocal chords must vibrate.

⁶ It was demonstrated by Frøkjær-Jensen et al. (1971) that in Danish, the aspirated plosives [p^h t^h k^h] have spread glottis, like the voiceless fricatives (e.g. [s f]), whereas the unaspirated plosives [b d ɡ] are voiceless but do not have (widely) spread glottis.

⁷ When the Sonority Syllable Model is applied to a specific language, some of its circles are typically non-active, i.e. they do not predict any specific ordering of segments in that language. This is no violation of the logic of the model: in an implication chain where ‘A implies B implies C’ it is, by logical necessity, also the case that ‘A implies C’ – this corresponds to the removal of a circle in the model. Furthermore, different circles can be non-active in either margin of the syllable.

⁸ The Sonority Syllable Model also predicts that marginal segments in absolute initial and final position have [spread glottis], for example *st-*, *ts-*; *-st*, *-ts*. This accounts for apparent violations of the sonority hierarchy and is testable by observing the glottis (Basbøll & Lambertsen 2014).

⁹ Danish /b d g/ are realized as voiceless, /p k/ as aspirated, and /t/ as affricated. Vowel qualities are normalized in accordance with Basbøll (2005) and Grønnum (2007).

¹⁰ There is no natural way to pronounce a Danish word like *hårdere* ‘harder’ with any contoids (like [ɤ]) intervening in the vocoidal string after [h].

¹¹ “Dano-Norwegian” in Haugen’s (1976) terminology.

¹² OLAM is a coding and analyses system for Danish developed by Claus Lambertsen, Hans Basbøll & Thomas O. Madsen.

¹³ The numbers in Table 6 are based on Table 3-5. Everything marked in the transcription as a segment is counted as a segment (e.g. final [h] even if it could be regarded as a breath). Superscript [ʰ] does not count as a segment, and the same goes for length, stress and *stød* as well as nasalization; identical segments (as in [ˈm:m:]) are each counted as a segment. The standard forms are very distinct (e.g. *skraldespanden* ‘bin’ without schwa-assimilation). In for example *kage* ‘cake’ [ˈkʰæ:(j)ə] “[j]” counts as a half a segment. Where there are two alternative forms, we count the mean.

¹⁴ The numbers in Table 7 are based on Table 3-5. For instance ['m:m:] counts as one syllable because there is only one stress and [m] is not an inherently syllabic segment. ['se:e:] is counted as two syllables due the two e's (because they are syllabic per definition), and a sequence [eæ] or the like is counted as two syllables. If syllabicity (by schwa-assimilation) is marked by "[-]", it counts as a syllable.

¹⁵ The numbers in Table 8 are based on Table 3-5. Here we count the number of different sonority types in the single word. Stød is a prosodic feature and therefore not taken into account here. The transcriptions are followed closely, i.e. a final stop which is transcribed as unaspirated and one which is transcribed as affricated or aspirated, are counted differently. The consonantal [ɣ] is counted as a voiced fricative. Nasals and [l] belong to the same sonority type and vocoids include vowels [ø], [w] etc.

¹⁶ The numbers in Table 9 are based on Table 3-5. We count one segment at a time, divided into sonority types. For example *blød* 'soft' [' ɸlødø?]: from [b] to [l] is one sonority rise, from [l] to [ø] likewise. The remaining segments are vocoids and the total number for *blød* is therefore 2.

Bibliographical References

- Basbøll Hans 1973. Notes on Danish Consonant Combinations. *Annual Reports of the Institute of Phonetics, University of Copenhagen* 7. 103-142.
- Basbøll Hans 2005. *The Phonology of Danish* (Series Phonology of the World's Languages). Oxford: Oxford University Press.
- Basbøll Hans 2012. Monosyllables and prosody: the Sonority Syllable Model meets the word. In Stolz Thomas, Nau Nicole & Cornelia Stroh (eds.), *Monosyllables: From Phonology to Typology*. Berlin: Akademie Verlag. 13-41.
- Basbøll Hans, Kjærbæk Laila, Lambertsen Claus & Ditte Boeg Thomsen 2012a. From Sound to Words: Discovering sound structure in Danish children's first words. In Henrichsen Peter Juel & Jan Heegård (eds.), *Speech in Action. Proceedings of the 1st SJUSK Conference on Contemporary Speech Habits. Copenhagen Studies in Language* 42. 195-217.
- Basbøll Hans, Kjærbæk Laila, Lambertsen Claus & Ditte Boeg Thomsen 2012b. Fra lyd til ord: En afdækning af lydstrukturen i danske børns første ord. *Danske Talesprog* 12. 182-207.
- Basbøll Hans & Claus Lambertsen 2014. La syllabe de sonorité rencontre le mot français: Aspects de morphonotaxe. In Durand Jacques, Kristoffersen Gjert & Bernard Laks (eds.), *La phonologie du français: normes, périphéries, modélisation*. Paris: Presses Universitaires de Paris Ouest. 83-99.
- Bleses Dorte, Basbøll Hans & Werner Vach 2011. Is Danish difficult to acquire? Evidence from Nordic past tense studies. *Language and Cognitive Processes* 26. 1193-1231.
- Bleses Dorte, Vach Werner, Wehberg Sonja, Thomsen Pia, Madsen Thomas O. & Hans Basbøll 2008. Early vocabulary development in Danish and other languages: a CDI-based comparison. *Journal of Child Language* 35. 651-669.

- Elbers Loekie & Josi Ton 1985. Play pen monologues: the interplay of words and babbles in the first words period. *Journal of Child Language* 12. 551-565.
- Elsen Hilke 1996. Two routes to language: stylistic variation. *First Language* 16. 141-158.
- Ferguson Charles A., Menn Lisa & Carol Stoel-Gammon (eds.) 1992. *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press.
- Frøkjær-Jensen Børge, Ludvigsen Carl & Jørgen Rischel 1971. A glottographic study of some Danish consonants. In Hammerich Louis L., Jakobson Roman & Eberhard Zwirner (eds.). *Form and Substance: Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen*. Copenhagen: Akademisk Forlag. 123-140.
- Jacobsen Henrik Galberg, Bleses Dorthe, Madsen Thomas O. & Pia Thomsen (eds.) 2003. *Take Danish for Instance – Linguistic Studies in Honour of Hans Basbøll*. Odense: University Press of Southern Denmark.
- Grønnum Nina 2003. Why are the Danes so hard to understand? In Jacobsen Henrik Galberg, Bleses Dorthe, Madsen Thomas O. & Pia Thomsen (eds.), *Take Danish for Instance – Linguistic Studies in Honour of Hans Basbøll*. Odense: University Press of Southern Denmark. 119-130.
- Grønnum Nina 2005. *Fonetik & Fonologi. Almen og dansk*. Copenhagen: Akademisk Forlag. 3d edn. [1st edn. 1998, 2d edn. 2001].
- Grønnum Nina 2007. *Rødgrød med fløde. En lille bog om dansk fonetik*. Copenhagen: Akademisk Forlag.
- Grønnum Nina 2008. Dansk udtale. Presentation in Dansk Sprognævn's Repræsentantskab [the Council of the official Danish language committee], October 3rd, 2008.
- Haugen Einar 1976. *The Scandinavian Languages*. London: Faber & Faber.
- Kent Raymond D. 1992. The biology of phonological development. In Ferguson Charles A., Menn Lise & Carol Stoel-Gammon (eds.). *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press. 65-87.
- Kjærbæk Laila 2013. *Tilegnelse af bøjningsmorfologi. En undersøgelse af substantivernes pluralisbøjning hos normaltudviklede danske børn i alderen 1-10 år*. Ph.D. dissertation, University of Southern Denmark, Odense.
- Ladefoged Peter 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Leopold Werner F. 1939. *Speech Development of a Bilingual Child: A Linguist's Record, vol. I: Vocabulary Growth in the First Two Years*. Evanston, IL: Northwestern University Press.
- MacWhinney Brian 2000a. *The CHILDES Project: Tools for Analyzing Talk, vol. 1: Transcription Format and Programs* (3rd edn.). Mahwah, NJ: Lawrence Erlbaum.
- MacWhinney Brian. 2000b. *The CHILDES Project: Tools for Analyzing Talk, vol. 2: The Database* (3rd edn.). Mahwah, NJ: Lawrence Erlbaum.
- Maekawa Junko & Holly L. Storkel 2006. Individual differences in the influence of phonological characteristics on expressive vocabulary development by young children. *Journal of Child Language* 33. 439-459.

- Oller D. Kimbrough 2000. *The Emergence of the Speech Capacity*. Mahwah: Lawrence Erlbaum Publishers.
- Oller D. Kimbrough & Eilers Rebecca E. 1988. The role of audition in babbling. *Child Development* 59. 441-449.
- Pike Kenneth L. 1943. *Phonetics: A Critical Analysis of Phonetic Theory and a Technic for the Practical Description of Sounds*. Ann Arbor: University of Michigan Press.
- Rischel Jørgen 2003. The Danish syllable as a national heritage. In Jacobsen Henrik Galberg, Bleses Dorte, Madsen Thomas O. & Pia Thomsen (eds.), *Take Danish for Instance – Linguistic Studies in Honour of Hans Basbøll*. Odense: University Press of Southern Denmark. 273-282.
- Smith Neilson V. 1973. *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.
- Sosa Anna V. & Carol Stoel-Gammon 2006. Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language* 33. 31-50.
- Stoel-Gammon Carol 1998. Sounds and words in early language acquisition. The relationship between lexical and phonological development. In Rhea Paul, Warren Steven F. & Mark E. Fey (eds.), *Exploring the Speech-Language Connection*. Baltimore: Paul H. Brookes Publishing Co. 25-52.
- Vihman Marilyn M. 1981. Phonology and the development of the lexicon: evidence from children's errors. *Journal of Child Language* 8. 239-264.
- Vihman Marilyn M. 1992. Early syllables and the construction of phonology. In Ferguson Charles A., Menn Lise & Carol Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press. 393-422
- Vihman Marilyn M. 1996. *Early Phonological Development. The Origins of Language in the Child*. Oxford: Blackwell Publishers.
- Vihman Marilyn M. & William Croft 2013. Phonological development: toward a "radical" templatic phonology. In Vihman Marilyn M. & Tamar Keren-Portnoy (eds.), *The Emergence of Phonology: Whole-word Approaches and Cross-linguistic Evidence*. Cambridge: Cambridge University Press. 17-57.
- Vihman Marilyn M. & Lorraine McCune 1994. When is a word a word? *Journal of Child Language* 21. 517-542.

Phonotactic processing and morpheme boundaries: word-final /Cst/ clusters in German

Chiara Celata, Katharina Korecky-Kröll, Irene Ricci & Wolfgang U. Dressler

Two experiments showed that, when processing consonant clusters of their native language, the speakers are sensitive to the presence of a morphemic boundary at some point of the cluster. In the first experiment, words with tri-consonantal final cluster had to be produced by inserting an /i/ in the cluster; the clusters containing a morpheme boundary were more frequently split according to the position of the boundary while the corresponding clusters included in a mono-morphemic word were split according to different options. In the second experiment, the subjects had to detect tri-consonantal clusters while a series of German words was auditorily presented; clusters included in monomorphemic words were easier to detect than homophonous clusters which spanned a morpheme boundary. The impact of morphonotactics on cluster processing was stronger on adolescents than on adults in Experiment 1. The results thus indicated that in offline and online processing of phonotactic structures, morphological information can be available and impact on the recognition and manipulation of clusters.

1. Introduction

Language-specific phonotactic patterns are known to affect multiple aspects of language processing, from segmentation of the speech stream (e.g. Saffran *et al.* 1996) to word learning (e.g. Storkel 2001), from phonological acquisition (e.g. Freedman & Barlow 2012) to morphological decomposition (e.g. Hay 2003). Phonotactics is exploited to locate word boundaries (e.g. McQueen 1998), but the speakers are also sensitive to differences between inter- and intra-word phonotactics (Mattys *et al.* 1999, Hay 2003). For example, words like *inhumane* in English are more likely to be decomposed than words like *insincere*, because the /nh/ transition across the morpheme boundary is not found word-internally in that language, while the /ns/ transition is well attested (e.g. *fancy*, *tinsel*) (Hay 2003: 16). Phonotactics is therefore also exploited to locate morpheme boundaries.

The phonotactics of morpheme concatenation is referred to as morphonotactics by Dressler & Dziubalska-Kołaczyk (2006), Dressler *et al.* (2010). Some studies observed that sequences occurring only or by default across a morpheme boundary (thus being

prototypical morphonotactic sequences) are acquired earlier and faster than lexical (or purely phonotactic) clusters by typically developing children in Lithuanian (Kamandulyte 2006) and Polish (Zydorowicz 2010); English children with grammatical-specific SLI perform better in regular past tense verbs when verbs contain clusters that are intra-morphemically legal than when they contain clusters that are only cross-morphemically legal (Marshall & van der Lely 2006). Consequently, it is claimed that morphology helps phonological acquisition: due to the interplay of phonotactics with morphotactics, morphonotactic sequences facilitate morphological and lexical acquisition. A facilitatory effect of morphological boundaries on phonological acquisition is attested in the recent literature also for word-final consonants in English: word-final sibilant fricatives are longer in plurals (such as *toes*) compared to monomorphemic words (such as *nose*) in adult speech, and children from 2 years of age onwards appear to be able to exploit such cues in fricative acquisition and to correctly reproduce the adult speech pattern (Song *et al.* 2013; see also Plag *et al.* 2015 for further evidence of morpho-phonetic effects in the realization of morphemic and non-morphemic word-final /s/ in a conversational English corpus). However, Freiburger (2014) investigated the early spontaneous speech data of three children acquiring Austrian German and did not find an advantage of morphonotactic over phonotactic clusters. Likewise, Korecky-Kröll & Dressler (submitted) investigated 3-year old children's spontaneous production of German consonant clusters and found no difference between morphonotactic and phonotactic clusters (although with different results according to the children's socio-economic status). Results on acquisition are thus less clear for German than for the other languages investigated.

Moreover, according to the Strong Morphonotactic Hypothesis, phonotactics is assumed to help morphological decomposition in adult processing too: if a certain sequence occurs only or by default over a morpheme boundary, it should be processed faster and more accurately than a comparable ordinary phonotactic sequence (Korecky-Kröll *et al.* 2014). In an experiment on German, visual targets corresponding to one phoneme (i.e., /t/) or two-phoneme sequences (i.e., /st/, /an/) were detected faster when separated from the word stem by a morphological boundary (e.g. *lobst* 'you praise', 2nd sg. present (*lob+st*); *bravste* 'bravest', superlative (*brav+ste*)) than when they were part of the word stem (e.g. *Obst* 'fruit'; *Bürste* 'brush'). This result was interpreted as a supportive effect of morphology on the speed with which individual phonemes or two-pho-

neme phonotactic sequences are processed.

If, however, morphology impacts the processing of clusters, then its effects should be visible also in the case of those sequences that straddle a morphological boundary. In particular, we may expect different processing for pairs of homophonous phonotactic sequences that may or may not occur across a morpheme boundary in the lexicon. These are called non-exclusively morphonotactic sequences in Dressler & Dziubalska-Kołaczyk (2006). For example, word-final /nst/ in German may appear in monomorphemic words (e.g. *Wanst*, ‘paunch’) as well as in the concatenation of the 2nd sg. verbal morpheme in /n/-final roots (e.g. *kannst*, ‘(you) can’). In this paper, clusters such as /nst/ in *Wanst* will be referred to as phonotactic clusters and clusters such as /nst/ in *kannst* as morphonotactic clusters.

Non-exclusively morphonotactic clusters can be preferentially phonotactic or morphonotactic, according to a gradual scale of preferences. The preference of a cluster to appear as phonotactic or morphonotactic in a language may be measured in terms of number of items in the lexicon (word types) or of the cumulated frequency of all lexical items in a corpus (number of tokens).

In a computational study of German (mor)phonotactic clusters (Calderone *et al.* 2014), a corpus-based model of neural network activation trained on a German corpus was used to ascertain if word-final German morphonotactic clusters produced a different representation from their purely phonotactic counterparts. The computational model, called PHACTS (Calderone & Celata 2012, Celata *et al.* 2011), simulated the formation of phonotactic knowledge in the mind of a German speaker who was exposed to a stream of phonological words by gradually developing a knowledge representation of the statistical regularities that shape the phonotactics of that language. Starting from the statistical knowledge that were deducible from the German CELEX database (Baayen *et al.* 1995), the model generated activation-based representations of full lexical forms and, crucially for our purposes, of the clusters included in those specific lexical forms. Experimental clusters were the non-exclusively morphonotactic sequence /nst/ and the exclusively morphonotactic sequences /mst/ and /xst/. The results showed that within-group variability in the activation value for the /nst/ clusters turned out to be significantly higher than for both /mst/ and /xst/ clusters. The results were interpreted as evidence in support of different phonotactic representations for the morphonotactic compared to the phonotactic /nst/ clusters. Additionally, the study suggested that the representation of morphonotactic clusters could emerge from corpus

information about the phonotactics of the language (including the co-occurrence of the clusters with specific vowels and consonant-vowel sequences, the relative frequency of vowel-consonant chunks and of phonological neighbors).

Do the speakers implicitly exploit such “proliferation of representations and extensive cross-activation” (Libben 2010: 12) when processing the consonantal clusters of mono- and bimorphemic words?

The present study aims at answering this question by testing the German speakers’ productive and perceptive reactions to tri-consonantal clusters with different (mor)phonotactic status. It therefore represents a psycholinguistic continuation of Calderone *et al.*’s (2014) study, from which it takes direct inspiration for the selection of the experimental materials. At the same time, this study is a continuation of Korecky-Kröll *et al.*’s (2014) study, inasmuch as behavioral data are measured for the processing of morphonotactic sequences that are larger than the morpheme, and specifically, for tri-consonantal clusters in which C1 is part of the lexical root and C2C3 coincide with an inflectional morpheme. In this study, we run two distinct experiments: a ‘split cluster’ game test, and a sequence monitoring task. The two experiments were aimed at evaluating the speakers’ reactions to the presence vs. absence of a morphological boundary during the phonological processing of the clusters. Given that the phonotactic and morphonotactic clusters that were compared in the experiments were homophonous (i.e., they were composed of the same segments in the same relative order), we postulated that any difference in subjects’ reactions had to be traced back to different phonological representations for homophonous phonotactic and morphonotactic sequences.

Concerning the developmental issue of whether and to what extent morphonotactic knowledge help in morphological and phonological acquisition, this study tests the existence of age-related effects in processing. To that aim, we opted for testing two different groups of subjects, one of adults and another of adolescent speakers. Given the intrinsic difficulty of the selected experimental procedures, it proved impossible to include children in the design. The hypothesis was that, if homophonous phonotactic and morphonotactic clusters are differently processed, then this difference should be more visible in the production and perception of adolescents, who are still in the process of increasing the size and complexity of their lexicon and grammar, than in the adults’ response.

2. Experiment 1: 'Split cluster' test

Experiment 1 was aimed at testing whether native speakers of German are sensitive to the presence of a morphological boundary within a bimorphemic cluster, when asked to produce a phonologically modified version of it. The experiment was conceived of as a language game in which the participants had to split the consonant clusters of monosyllabic words by inserting an /i/ vowel between the consonants of the cluster. In 80% of the cases, these were bi-consonantal clusters and the subjects had to insert the vowel between the first and the second consonant. In the remaining 20%, however, the items contained tri-consonantal /Cst/ clusters and the subjects had to decide whether inserting the vowel either between the C and the /s/ or between the /s/ or the /t/. For example, for a word such as *Wulst* 'bulge', the participants could either produce /'vulist/ or /'vulsit/. The experimental words included both words ending in a phonotactic cluster, such as *Wulst*, and words ending in a morphonotactic cluster (with /st/ as the second morpheme of the word, e.g. *lobst* 'you praise'). Our prediction was that for morphonotactic clusters the /Cist/ response was more frequent and/or easy than for the phonotactic ones. The reason of such expectation is that splitting the /Cst/ word-final cluster as /Cist/ corresponds to segmenting the original word into its morphological constituents (root + morphological ending; e.g. *lob-st* > *lob-i-st*), which is not the case for words containing a phonotactic cluster (where no morphological boundary is found).

2.1. Participants

Two groups of native Austrian German speakers participated in the study. The first was a group of 38 adults (age 29-52, 20 females, 18 males). The second group was composed of 26 adolescents (age 11-15, 17 girls, 9 boys) attending a high school (AHS) in Vienna. None of them reported any speech or hearing disability.

2.2. Materials

The experimental items were 14 German monosyllabic words ending in a /Cst/ cluster (in one case a /CCst/ cluster). Half of them contained a phonotactic cluster and the other half a morphonotactic one. Eight items ended in /nst/, two in /pst/, two in /lst/ and two in /ŋkst/ (an Austrian variant of /ŋst/). The items are listed in Table 1. The two classes (phonotactic vs. morphonotactic clusters) were balanced for average word frequency (calculated as the number of occurrences in the 500-million-word *Leipzig Deutscher Wortschatz Online* database).

Table 1. German monosyllabic stimuli used as experimental items in Experiment 1. Verbal forms correspond to 2nd sg present.

CLUSTER TYPE	PHONOTACTIC ITEMS		FREQUENCY	MORPHONOTACTIC ITEMS		FREQUENCY
/nst/	<i>Dunst</i>	'mist'	386	<i>kennst</i>	'know'	102
	<i>Wanst</i>	'paunch'	5	<i>kannst</i>	'can'	1661
	<i>Kunst</i>	'art'	6325	<i>meinst</i>	'mean'	102
	<i>Gunst</i>	'goodwill'	1388	<i>nennst</i>	'call'	21
/pst/	<i>Obst</i>	'fruits'	869	<i>lobst</i>	'praise'	94
/lst/	<i>Wulst</i>	'bulge'	20	<i>weilst</i>	'stay'	45
/ŋkst/	<i>Angst</i>	'fear'	17370	<i>längst</i>	'long ago'	19756
<i>average</i>			3766			3112

Fifty-six monosyllabic German words containing a bi-consonantal cluster served as fillers. Ten were /ns/-, /ps/- and /ls/-ending words (called 'CS fillers'); 6 of them were monomorphemic, while 4 contained a morphonotactic cluster. An additional group of 10 /st/-ending words ('ST fillers') also included 6 monomorphemic and 4 morphonotactic clusters. Another 10 words ended in a variety of bi-consonantal clusters (/ld, rf, nts, rx, rb, nt, rf, mp/; these were called 'CC fillers'). Finally, 26 words with a bi-consonantal cluster in initial position were included ('w_initial fillers').

Two experimental lists were created by pseudo-randomizing the 70 total items (14 experimental items and 56 fillers). The randomization was controlled in order to avoid the occurrence of a CS or ST filler before an experimental item; two consecutive experimental items were also banned. The participants were assigned one of the two lists in equal proportion.

Ten additional monosyllabic words ending (N = 6) or beginning (N = 4) with a biconsonantal cluster (word-final /st lf lm xs/ and word-initial /kr dr sl sn/) served for the training session.

2.3. Procedure

Experiment 1 was an offline test implemented with Presentation® on a Windows PC. The participants were requested to take part in a language game. They had to assume that, in a Martian language very similar to German, any word can be modified by inserting a /i/ vowel between the consonants of a cluster, to create diminutive and attenuative forms (e.g. *Hund* 'dog' → /'hunit/ 'small dog' or *fragen* 'to ask' → /fi'ra:gen/ 'to ask a few questions'). They were then presented with one of the two experimental lists and asked to insert /i/ as fast as possible whenever they found a C cluster in the word stimulus.

The stimuli were presented one at a time. Each trial began with a fixation stimulus (*****) of 500 ms, after which the word stimulus was presented, both

visually (capitalized, in the center of the screen) and orally (on headphones). The duration of the visual stimulus was equivalent to the duration of the oral stimulus; each individual stimulus had a different duration, which could vary between 531 and 617 ms. The subjects were instructed to repeat the word with an /i/ between the two consonants of the cluster as soon as they could (although any response given within 3500 ms from the offset of the stimulus was taken as valid and the reaction times were not measured). They were not told that the experiment included a minority of tri-consonantal clusters.

Before the experiment, a training session consisting of 10 items with bi-consonantal clusters was presented.

2.4. Analysis

A total of 532 (adult group) and 364 (adolescent group) responses for the experimental items was analyzed. Cross-tabulations of response type ('IST' response, 'SIT' response, missing response) by stimulus type (phonotactic, morphonotactic) were realized, separately for the two age groups. The two lists did not differ for the percentage number of missing responses, nor for the overall distribution of the response types; the data of the two lists were therefore collapsed for the purposes of the present analysis.

Two-sided Pearson's Chi Square tests were run to evaluate the presence of significant differences in the distribution of the response types across groups and stimulus types.

2.5. Results

The results of the cross-tabulations are shown in Table 2. Overall, the 'IST' response was more frequent than the 'SIT' response (59% vs. 33%). The group of adults produced a lower percentage of missing responses compared to the adolescents (4% vs. 12%); this indicated that the task turned out to be not as easy for the younger participants as for the older ones.

For the group of the adult participants, no significant effect of stimulus type was found on response type (χ^2 (5, $N = 532$) = 1.858, $p > .05$). This meant that the splitting strategy did not change significantly as a function of the type of the cluster included in the item; the proportion of missing responses (i.e., the failure to split the cluster) was also very similar for the phonotactic and the morphonotactic items (though the former turned out to be on average slightly more difficult to split than the latter).

By contrast, in the case of the adolescents, the test revealed the presence of an uneven distribution of the different types of response. In particular, the phonotactic clusters were significantly more difficult to deal with compared to the morphonotactic ones (17.6% vs. 7.1% of missing responses, respectively) and less frequently split according to the 'SIT' response (53.3% vs. 64.3%). Such differences were statistically significant (χ^2 (5, $N = 364$) = 9.901, $p < .01$).

No additional effect of stimulus or consonantal context (/nst pst lst ŋkst/) on the proportion of response types was found (stimulus: adults $\chi^2(20, 532) = 28.729$, $p > .05$; adolescents $\chi^2(20, 364) = 35.310$, $p > .05$; context: adults $\chi^2(11, 532) = 10.315$, $p > .05$; adolescents $\chi^2(11, 364) = 8.288$, $p > .05$).

Table 2. Frequency and percentage of ‘IST’, ‘SIT’ and missing responses as a function of age group and stimulus type in Experiment 1.

		ADULTS		ADOLESCENTS	
		N	%	N	%
Phonotactic	missing	14	5,3	32	17,6
	‘IST’	155	58,3	97	53,3
	‘SIT’	97	36,5	53	29,1
	Total	266	100	182	100
Morphonotactic	missing	8	3	13	7,1
	‘IST’	163	61,3	117	64,3
	‘SIT’	95	35,7	52	28,6
	Total	266	100	182	100

2.6. Discussion

Experiment 1 aimed at testing whether native speakers of German are sensitive to the presence of a morphological boundary within a tri-consonantal cluster when asked to disrupt the consonantal sequence by inserting an /i/ vowel internally. In particular, it was hypothesized that for those items where a morphological boundary occurs between the first and the second consonant, the vowel was inserted precisely in that position significantly more often than for items with no internal morphological boundary. It was also hypothesized that an age effect could emerge, with younger speakers more sensitive to the difference between phonotactic and morphonotactic clusters than older speakers; to test that hypothesis, the performance of adult native speakers was compared with that of a group of adolescents.

The results showed that, overall, the subjects opted for the ‘IST’ response more frequently than for the ‘SIT’ response for both phonotactic and morphonotactic items. This result can be explained by making reference to general principles of phonotactic markedness: all other things being equal, a sequence like /CVCist/ is phonotactically less marked than a sequence like /CVCsit/ (with a word-internal obstruent-sibilant cluster). Language-specific preferences may

also play a role here: word-final /-ist/ is more frequent than word-final /-sit/ in German.

A differential effect of morphological congruency in the case of the 'IST' response type was expected to hold for words with morphonotactic clusters only. In this respect, the results pointed towards the existence of a significant difference between the performance of adult participants and that of the adolescents. Only the latter, and not the former, treated morphonotactic clusters differently from their phonotactic counterparts. This was evident in two respects: first, morphonotactic clusters were easier to split; second, they were comparatively more often split according to the 'IST' option, that is, by separating the C-final root from the /st/ inflectional ending. These data can therefore be interpreted as a heavier reliance of adolescents (compared to adults) on the morphonotactic status of clusters, that is, items containing a morphonotactic cluster were easier to be divided into two parts given that they are composed of two morphemes. This is compatible with the general hypothesis that knowledge of the morphonotactic status of clusters improves the adolescents' processing of the stimuli.

To sum up, Experiment 1 demonstrated that, in an offline test requesting to produce a modified version of homophonous mono- and bi-morphemic clusters, the youngest speakers used implicit lexical and morphological information as a possible cue for the accomplishment of the phonological production task. This suggests that morpheme boundaries may indeed count as relevant information cues in differentiating homophonous clusters, as assumed under the Strong Morphonotactic Hypothesis (see above).

At the same time, this opens two further questions. The first question is whether morpheme boundaries also play some role in online, pre-lexical processing of homophonous clusters. If they do, a second question follows, and specifically, whether they play a facilitatory effect (as in the case of the offline lexical production task of Experiment 1) or an inhibitory one (as one could assume, starting from the observation that the presence of a morpheme boundary within a consonant sequence represents an additional information to process and, as such, could imply an additional processing cost).

To address these issues, Experiment 2 was conceived to investigate whether the effect of morphonotactics reported for the adolescents in the 'split cluster' test was also present in an online perceptual experiment focusing on pre-lexical processing.

3. Experiment 2: Fragment monitoring

Experiment 2 involved a slightly modified version of the sequence monitoring paradigm of Frauenfelder & Kearns (1996). This task is mostly known in the form of syllable monitoring as it was extensively used to investigate the role of syllables in speech processing and lexical access (since Mehler *et al.* 1981). In sequence monitoring tasks, subjects are presented with targets that are either congruent or incongruent with a linguistic unit in the target-bearing item. They have to decide whether the target is part of the spoken input that is presented subsequently. Congruent targets yield faster detection latencies, compared to incongruent ones. In the experiment presented here, the targets were consonant sequences such as /nst/, /pst/, /lst/ and /kst/. Consistently with the standard variant of the sequence monitoring task, the targets were presented visually, and spoken carrier words had to be monitored for the occurrence of these targets. The targets occupied the final position in monosyllabic words or the internal position in disyllabic words. As in the preceding experiment, the clusters were phonotactic in half of the cases, morphonotactic in the other half (with /st/ or /sta/ as the second morpheme, or second and third morphemes, of the word, in mono- and disyllables respectively). Our expectation was that, if morpheme boundaries play a role in the monitoring task, morphonotactic clusters are longer and less accurate to detect than phonotactic clusters, because only the former, and not the latter, are incongruent with linguistic units (in this case, morphemes) of the target-bearing item.

3.1. Participants

As in the case of the previous experiment, two groups of Austrian German native speakers participated in Experiment 2. The first group was composed of 41 adults (age 25-59, 24 females, 17 males). The second group was composed of 28 adolescents, aged 12-16, 16 girls and 12 boys, attending the same Viennese high school as those of Experiment 1. None of them reported any speech or hearing disability.

3.2. Materials and procedure

The experimental items were 30 German words containing a cluster of three (in a few cases, four) consonants in either word-final or word-medial position. The clusters were /nst/, /lst/, /pst/ and /kst/; all were phonotactic in half of the cases, morphonotactic in the other half. Eight words were disyllables with the cluster in internal position; the remaining 22 words were monosyllables with the cluster in the final position (see Table 3). The composition of the two groups

(phonotactic vs. morphonotactic items) was balanced for as many parameters as possible: word length (number of syllables and of phonemes), position of the cluster (word-medial and word-final), type of cluster (/n/, /l/, /p/ and /k/ followed by /st/) and average frequency, calculated as number of occurrences in the *Leipzig Deutscher Wortschatz Online*. However, while average frequency for the phonotactic and the morphonotactic items was very similar for /nst/ and /pst/ items, it was not for /lst/ and /ɣst/ items; in the former case, morphonotactic items were more than twice as frequent as the phonotactic ones, while in the latter, they were more than 10 times less. Finally, all phonotactic clusters were immediately preceded by a short vowel, except *Papst* ‘pope’ and *Obst* ‘fruits’, where the vowel (which may be long or short in Austrian German) was realized as long. All morphonotactic clusters were immediately preceded by a short vowel, except *schonst* ‘you spare’, *sagst* ‘you say’ and *gibst* ‘you give’ (which may have a long or short vowel in Austrian German), where the vowel was long.

The duration of the cluster of each experimental item and the duration of the speech portion preceding the cluster were measured. The clusters’ average duration for phonotactic and morphonotactic items was approximately the same. In the case of the preceding segments, there was a difference of about 15 ms (with phonotactic clusters shorter than morphonotactic ones). An independent sample t-test with cluster duration and pre-cluster duration as dependent variables and cluster type (phonotactic, morphonotactic) as grouping factor showed that the two groups of items did not differ significantly (cluster duration: $t(28) = -0.173$, $p > .05$; pre-cluster duration: $t(28) = -0.973$, $p > .05$). However, as the data in Table 3 shows, there was much variation according to cluster type. Particularly in the case of pre-cluster duration, the difference between phonotactic and morphonotactic /nst/ items was one of 64 ms (morphonotactic items having shorter pre-cluster durations), while for /lst/ items was of about 35 ms (morphonotactic items having longer pre-cluster durations); also in the case of /kst/ clusters, pre-cluster duration in morphonotactic was on average 30 ms longer than in phonotactic items.

There is currently big debate about whether seemingly homophonous segments that do, or do not, represent morphemes are phonetically identical or different. A growing amount of evidence shows indeed that subtle durational differences may cue functional distinctions between morphemic and non-morphemic segments, or between different morphemes (e.g. Plag *et al.* 2015 for a large-scale account of word-final /s/ in English). The variability found in our data for clusters and for their phonological environments could in principle hide a morpho-

phonetic effect of this kind. However, our word set is too small and does not allow generalizations. We therefore limit our phonetic analysis to the observations above; since between-group durational variability could in principle affect the subjects' response, we included cluster and pre-cluster durations as predictors in the statistical model.

Table 3. Experimental stimuli used in Experiment 2. “CDur” = cluster duration (in ms); “PCDur” = pre-cluster duration (in ms); “Freq” = word frequency.

CLUSTER TYPE	PHONOTACTIC ITEMS				MORPHONOTACTIC ITEMS			
		CDUR	PCDUR	FREQ		CDUR	PCDUR	FREQ
/nst/	sonst	353	201	12691	schonst	355	248	6
	Wanst	396	177	22	kannst	415	146	29709
	Kunst	413	168	31593	kennst	409	147	21628
	Gunst	395	75	2715	nennst	400	177	155
	Monster	352	153	983	dünnste	275	89	62
/nst/ average		382	155	9601		371	161	10312
/pst/	Papst	271	164	16971	knipst	332	154	113
	Herbst	334	263	3484	plumpst	288	220	35
	Obst	330	145	6058	gibst	357	170	24740
/pst/ average		312	191	8837		326	181	8296
/lst/	Wulst	368	207	29	willst	374	155	3768
	Elster	440	73	99	hellster	441	147	4
	Polster	372	119	1431	vollste	419	208	104
/lst/ average		393	133	520		411	170	1292
/kst/	Text	344	113	17026	mixt	345	155	215
	Gangster	346	178	1313	längste	367	211	2423
	Angst	308	189	5737	denkst	327	159	694
	Axt	365	106	88	sagst	326	189	1069
/kst/ average		341	147	6041		341	179	441
clusters average		359	155	6683		362	171	5648

Seventy-five mono- and disyllabic German words served as fillers. There were two experimental lists. Each list was divided into 16 trials. There were 11 ‘go’ and 5 ‘no-go’ trials. Each ‘go’ trial began with the visual presentation of an orthographic string correspond-

ing to the clusters of the experiment: these were NST (for /nst/ clusters), LST (for /lst/ clusters), PST and BST (for /pst/ clusters), and XT, GST and KST (for /kst/ clusters). This target string appeared in the middle of the computer screen, in capitals and with an appropriate font dimension, and remained visible for 2 seconds. The stimuli were 4 to 8 words of which 1 or 2 were experimental phonotactic or morphonotactic items containing the phoneme sequence indicated by the visual string (e.g. in the case of the string LST, the items contained the sequence /lst/). The others were fillers and did not contain those phonemes (neither the sequence /lst/ nor the individual phonemes /l/, /s/ or /t/). The 'no-go' trials were identical to the 'go' trials except that they did not contain any word item matching the phonological sequence indicated by the target string; the initial strings in the 'no-go' trials were FST, BST, MS and PS. The 'no-go' trials were identical in the two experimental lists. However, the 'go' trials were differentiated across lists inasmuch as, for each target string, one list contained only the phonotactic item as matching word, the other list contained only the corresponding morphonotactic item. The proportion of phonotactic and morphonotactic items as matching words was balanced in the two lists. Each list was therefore composed of 90 words (75 fillers and 15 experimental items) subdivided in 16 trials. Each list appeared with two different orders of trial presentation (two randomizations), resulting in a total of four final lists. Each subject was assigned one of the four final lists.

The oral stimuli had been produced by a male native speaker of Standard Austrian German and recorded on an Edirol 4-Channel portable Recorder and an AKG CK91 microphone with a sampling rate of 44100Hz/16 bit. After the presentation of the target string, the relevant oral stimuli were consecutively presented on headphones. For each stimulus, the subjects had to press a 'yes' key (the green button on a Cedrus response pad) if they thought that the word contained the phonemes indicated by the target string or a 'no' key (the red button) if they thought differently. The time lag for the response was 2000 msec. The target string remained visible during the presentation of the associated stimuli. The inter-trial interval was 1000 msec.

Two trials of 6 and 4 oral stimuli preceded by, respectively, ST and NZ target strings served as training session.

3.3. Analysis

The dependent variables were response latencies and accuracy. As fixed factors we considered Age (adults vs. adolescents), Frequency (= the log transformed frequency of occurrence of each

experimental item in the reference corpus), ClusterDuration (in msec), PreClusterDuration (= the duration of each item stimulus from its onset to the onset of the target cluster; in msec), AntecedentRT (= the response latency of the preceding item in the experimental list; for latency analysis only), AntecedentAccuracy (= whether the preceding item yielded a correct response or an error; for accuracy analysis only), Status of the cluster (phonotactic vs. morphonotactic), and the interactions Age*Status of the cluster and Age*Frequency.

Mixed effects regressions were implemented in the packages lme4 of the software R (Bates *et al.* 2014). To bring the variation of random effects such as subject or item under statistical control, a mixed-effect model with subjects, items and cluster types as random effects was run. A factor was considered significant if its t-statistics yielded a t-value of greater than 2 (or less than -2) when included in the model, or its *p*-value was < .05; in addition, a likelihood ratio test compared the model including the Status of the cluster factor to a model without it; the test had to yield a *p*-value lower than 0.05 to show that the inclusion of the factor did significantly improve the fit of the model.

A total of 3780 responses for the adults and 2340 for the adolescents were analyzed for global accuracy, i.e. number of correct responses ('yes' button for matching words, 'no' button for non-matching words). The subsets of 630 responses for the adults and of 390 for the adolescents corresponding to the 30 experimental items were analyzed for accuracy (i.e., the number of correct identifications of the relevant fragment in the "go" trials). Of these, the correct responses (N = 556 for the adults; N = 308 for the adolescents) were also analyzed for reaction times.

3.4. Results

3.4.1. Accuracy

The analysis of global accuracy showed that, overall, the adolescents performed more errors than the adults (6% vs. 3.5%, respectively). The same result was found when we considered the subset of experimental items only: in this case, the percentage of missed identifications was 21% for the adolescents and 12% for the adults. As in the previous experiment, then, the younger speakers found the task relatively more demanding compared to the adults.

The data in Table 4 also reveal that adolescents made more errors on the morphonotactic items than on the phonotactic ones (respectively, more than 25% of the morphonotactic items were misidentified, compared to 16% of the phonotactic ones). The difference

was comparatively less evident in the case of the adults' performance (13% vs. 10%), which suggested that the two groups of subjects were differently affected by the (mor)phonotactic status of the cluster in the monitoring task.

Table 4. Accuracy in the identification of experimental items as a function of stimulus type and age range. C = correct identification, E = error (missed identification of the matching fragment).

			PHONOTACTIC	MORPHONOTACTIC	TOTAL
ADULTS	C	Frequency	283	273	556
		% within Phonot./Morphonot.	89,9	86,6	88,3
	E	Frequency	32	42	74
		% within Phonot./Morphonot.	10,1	13,4	11,7
	Total	Frequency	315	315	630
ADOLESCENTS	C	Frequency	163	145	308
		% within Phonot./Morphonot.	83,6	74,4	79
	E	Frequency	32	50	82
		% within Phonot./Morphonot.	16,4	25,6	21
	Total	Frequency	195	195	390

The mixed effects logistic regressions with accuracy as dependent variable and subjects, items and cluster types as random factors was aimed to verify whether the probabilities associated to incorrect detection of clusters were associated with changes in the status of the cluster (from phonotactic to morphonotactic) and/or in the age of the participants.

The results showed that the only significant factors turned out to be Age, with adult participants significantly more accurate than adolescents, and (log transformed) Frequency, with low frequency items more frequently hit by errors than high frequency items (see Table 5). Neither the Status of the cluster factor nor the Age*Status and Age*Frequency interactions showed up as statistically significant.

The observed differences in the average accuracy scores, therefore, could not be confirmed by the mixed effects logistic regression statistics.

Table 5. Fixed-effect coefficients and *p* values in the mixed-effects model fitted to accuracy (subjects, items and cluster types as random factors). The reference levels for the categorical predictors are the following: for Age it is ‘adolescents’, for Status of the cluster it is ‘morphonotactic’. All coefficients can be interpreted as changes relative to these reference levels.

COEFFICIENTS	ESTIMATE	STD. ERROR	Z VALUE	Pr(> z)
(Intercept)	0.126	1.5383	0.082	0.934
Ageadult	-0.822	0.275	-2.978	**
Frequency	-0.111	0.0378	-2.942	**
AntecedentAccuracy	0.277	0.710	0.391	0.695
PreClusterDuration	0.002	0.002	0.923	0.356
ClusterDuration	-0.003	0.003	-0.988	0.322
Statusphono	-0.368	0.268	-1.372	0.170
Ageadult*Statusphono	0.247	0.363	0.682	0.495
Ageadult*Frequency	0.222	0.267	0.685	0.345

3.4.2. Reaction times

In the mixed-effect model with subjects, items and cluster types as random factors and latencies as dependent variable, we wanted to verify whether the phonotactic or morphonotactic status of the cluster affected the subjects’ reaction times in the detection task, and whether this effect was modulated by the age factor. Figure 1 shows by means of boxplots that for both groups of subjects, the median lines are higher for the morphonotactic than for the phonotactic items; but there seems to be a bit more overlap (and more dispersion) in the adolescents’ performance than in the adults’ one.

The results of the mixed model are shown in Table 6. There were, as expected, significant effects of (log transformed) Frequency (higher frequency items faster than lower frequency items), AntecedentRT (the faster the antecedent item, the faster the target item), and PreClusterDuration (the longer the duration of the stimulus up to the cluster’s onset, the longer the time needed to give the response). The duration of the cluster, by contrast, did not affect the model. Age yielded a significant result, with adult participants reacting faster than adolescents (195 msec on average). Most importantly, our predicted explanatory variable Status of the cluster also yielded a significant result, showing that phonotactic clusters were detected on average 73 msec earlier than the corresponding morphonotactic ones. However, the interaction Age*Status was not significant, suggesting that the two groups of participants were equal in showing shorter latencies in the detection of the phonotactic clusters.

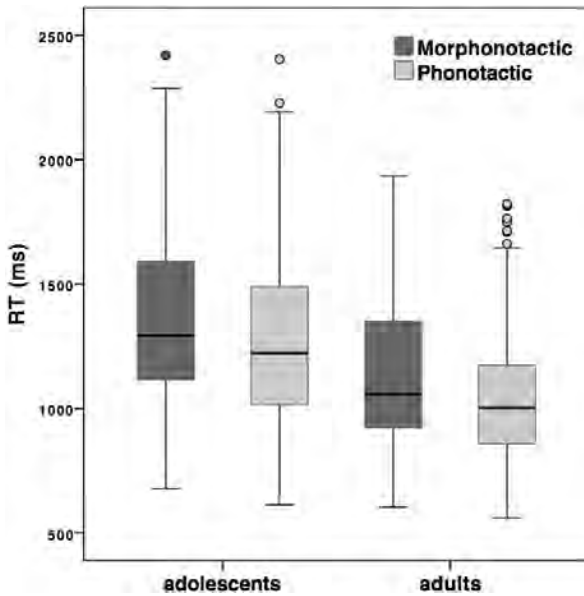


Figure 1. Relationship between age groups and the status of the cluster (phonotactic, morphonotactic). (Error bars represent confidence intervals of the individual categories.)

Another interesting result was related to the Age*Frequency interaction, which turned out to be very close to significance. In order to assess the explanatory contribution of item frequency to response latencies in the two groups of speakers, two linear regression models for the two subject groups were run. For the group of the adult speakers, the two variables turned out to be negatively and significantly correlated (Pearson correlation coefficient $r = -0.211$, $r^2 = 0.45$; $F(1,542) = 25,256$, $p < .0001$). This meant that the higher the word frequency, the shorter the time needed to detect the cluster in the word, as expected. For the group of the adolescents, by contrast, the two variables turned out not to be significantly correlated (Pearson's correlation coefficient $r = -0.105$, $r^2 = 0.16$; $F(1,295) = 2,655$, $p > .05$). Therefore, in the adolescents' performance, the response speed was in no way influenced by the fact that the target sequence could occur in high vs. low frequency words. These data suggested slightly different processing strategies in cluster detection for different age groups.

Table 6. Fixed-effect coefficients and *p* values in the mixed-effects model fitted to RT (subjects, items and cluster types as random factors). The reference levels for the categorical predictors are the following: for Age it is ‘adolescents’, for Status of the cluster it is ‘morphonotactic’. All coefficients can be interpreted as changes relative to these reference levels.

COEFFICIENTS	ESTIMATE	STD. ERROR	T VALUE
(Intercept)	1250.470	114.508	10.920
Ageadult	-194.843	49.280	-3.956
Frequency	-13.158	3.502	-3.757
AntecedentRT	0.225	0.038	5.961
PreClusterDuration	0.471	0.182	2.586
ClusterDuration	-0.357	0.239	-1.489
Statusphono	-73.464	27.484	-2.673
AgeGroupadult*Statusphono	8.013	33.742	0.237
AgeGroupadult*Frequency	-0.348	0.221	-1.989

We additionally checked the effect sizes of the explanatory variable by fitting a ‘null’ model that lacked the Status of the cluster predictor, and considered whether the difference between the likelihood of these two models (the full and the ‘null’ model) is significant. The results of the likelihood ratio test confirmed that the morphonotactic status of the clusters significantly affected reaction times ($\chi^2(2)=16.75$, $p=0.00023$).

3.5. Discussion

Experiment 2 was aimed at testing whether native speakers of German were sensitive to the presence of a morphological boundary within tri-consonantal clusters when asked to identify them as fast as possible in isolated words presented orally. In particular, it was hypothesized that, if morphological boundaries played a role in cluster detection, morphonotactic clusters were slower and less accurate, due to the fact that the first consonant pertains to the lexical morpheme composing the word while the second and third consonants form (part of) the inflectional morpheme. While Experiment 1 provided some evidence in support of the morphonotactic hypothesis by pointing to the facilitatory effect of morphological structure in cluster modification (production), Experiment 2 was expected to test the role of morphonotactics in pre-lexical phonological processing (perception). As in Experiment 1, it was also hypothesized that an age effect could emerge, with younger speakers more sensitive to the morphonotactic status of clusters than older speakers; in order to test that hypothesis,

the performance of adult native speakers was compared with that of a group of adolescents. Unlike Experiment 1, in which the participants had no temporal restriction for the execution of the task, Experiment 2 tested the online processing of clusters by constraining the temporal interval allowed for the response.

The first expectation, concerning the difference between phonotactic and morphonotactic clusters, was confirmed by the reaction time analysis, but not by the accuracy analysis. In addition, the effect that was found on latencies was an inhibitory one. Morphonotactic items turned out therefore to be slower, but not less accurate, to be detected than phonotactic items. The latency results can thus be interpreted in support of the morphonotactic hypothesis, inasmuch as the presence of a morphological boundary within consonant clusters turned out to affect the speakers' time of recognition of those clusters; since this effect was inhibitory, the result also support the view that, in this online perceptual experiment, morpheme boundaries represented an additional processing cost during the phonological processing of the items.

The second expectation, concerning different processing strategies by different age groups, was basically disconfirmed by the results of Experiment 2, although the data were not entirely clear. In the accuracy analysis, the Age*Status of the cluster interaction was not significant, though the average values pointed to an existing trend for adolescent speakers to produce more errors on the morphonotactic items than on the phonotactic. In the latency analysis, the interaction was equally insignificant, though the adolescents turned out to be overall insensitive to word frequencies. The latter point is worth of further comment.

In the adolescents' performance, the apparent independence of reaction times from lexical frequencies could suggest that, when performing the sequence monitoring task, these participants accessed the phonological representation of the cluster without necessarily relying on the lexical representation of the whole word. For them, lexical properties such as whole word frequency were indeed irrelevant. Then, if we consider once again that the sequence monitoring task is based on the assumption that congruent targets yield faster detection latencies than incongruent targets, and that in Experiment 2, consonantal sequences appearing in multimorphemic words were said to be incongruent with the morphemic structure of the word, the frequency result on adolescents would provide indirect support to a view of morphological processing in which individual morphemes are recognized early, and morphemic representations are activated prior to the activation of lexical entries (from Taft & Forster 1975 onwards, with different versions of the dual-route approach, particularly as far as

morpho-orthographic representations are concerned; e.g. Caramazza *et al.* 1988, Longtin *et al.* 2003, Taft & Ardasinski 2006).

4. General discussion and conclusions

By comparing the behavior of adults and adolescents in splitting, modifying and detecting tri-consonantal phonotactic and morphonotactic sequences of their native language, the two experimental studies presented here aimed to investigate the processing of phonotactic sequences that straddle a morphological boundary, in comparison to homophonous monomorphemic sequences.

Experiment 1 demonstrated that, when asked to disrupt a consonantal sequence and introduce an /i/ vowel at some point between two consonants of the cluster, part of the speakers took advantage of the information related to the morphological structure of the word, thus inserting the vowel preferentially at the point of the morphemic juncture; and in general, bimorphemic clusters were easier to split and modify than monomorphemic ones. This result provided support to the general claim that the speakers are sensitive to the phonotactic vs. morphonotactic status of homophonous consonant clusters, and may use this information in processing the phonological and morphonological characteristics of words. The fact that such sensitivity was found in adolescents only, not in adults, also pointed to age-related differences in (mor)phonotactic processing. As a matter of facts, it is well known that the acquisition of phonology starts earlier and is mastered earlier than the acquisition of morphology; as a consequence, the interaction between these two components necessarily changes in early childhood (Dressler & Dziubalska-Kolaczyk 2006; Bavin 2009). Additionally, bootstrapping between phonology and morphology is limited to early phases of acquisition (Weissenborn & Höhle 2000). However, our data from Experiment 1 only indirectly support this view, inasmuch as it is not entirely clear, first of all, whether and to what extent adolescents also differ from pre-pubescent children. Second, similar age-related differences in (mor)phonotactic processing were neither consistently nor significantly found in Experiment 2, contrary to the expectations.

Capitalizing on the suggestions coming from Experiment 1 that adolescents differed from adults in the way they exploited morphonotactic information in producing phonologically modified consonant clusters, Experiment 2 aimed at verifying whether the same sensitivity to the clusters' (mor)phonotactics was equally present in the

online pre-lexical processing of phonological sequences. The results of Experiment 2 were consistent with those of Experiment 1 inasmuch as phonotactic and morphonotactic clusters were processed differently (i.e., the former elicited shorter reaction times than the latter), but diverged to the extent that the response pattern was generalized to adults too. Adults and adolescents did differ, to the extent that the latter were overall less accurate and slower than the former; but this was indicative of a higher processing cost associated to the accomplishment of the experimental task for adolescents as compared to adults; age differences did not influence the way the speakers processed the contrast between morphonotactic and phonotactic cluster. Furthermore, while Experiment 1 showed that morphonotactic clusters were easier to split and produce (with the addition of an /i/ vowel) than phonotactic ones, Experiment 2 showed that morphonotactic clusters were slower to be detected than phonotactic ones (though eliciting an equal number of correct and incorrect responses).

Taken together, these results indicate that morphonotactics can be relevant in phonological processing, but its influence varies as a function of experimental task, and, in some experimental tasks, according to the age of the speakers. In online phonological processing, morphonotactic sequences may involve an additional processing cost compared to homophonous phonotactic sequences. This processing cost may be the same for adults and for younger speakers. Korecky-Kröll *et al.* (2014), also on German, demonstrated that adult native speakers actively exploit the potential boundary signaling function of clusters that result from morphological concatenation, by treating morphologically produced clusters differently than morpheme internal clusters. The results of Experiment 2 consistently showed that the two types of clusters can be processed differently by both adolescents and adults. The fact that in the study by Korecky-Kröll *et al.* (2014) the morphonotactic sequences were advantaged in processing, compared to the phonotactic ones, was due to the fact that those subjects were asked to detect phonemes or sequences that were either separated from the stem by a morphological boundary, or not; these sequences did not straddle a morphological boundary. In our Experiment 2, by contrast, morphonotactic clusters were incongruent with the monitoring target, which could explain their disadvantage in processing. Therefore, both the present study and the one of Korecky-Kröll *et al.* (2014) consistently support, from two different perspectives, the observation that the presence of a morphological boundary is relevant to the speakers in their processing of consonantal clusters of their native language.

The present findings are also consistent with the results of the psycho-computational study of Calderone *et al.* (2014), in which the phonological representation of morphotactic sequences and of homophonous phonotactic ones were derived from an artificial algorithm trained on German phonotactics and turned out to be significantly different, independently from any quantitative parameter of phonotactic probability or lexical frequency. In that paper it was hypothesized (Calderone *et al.* 2014: 69) that the early phases of lexical and morphological processing are characterized by automatic parsing attempts and the need to extract as much information as possible from language input, thus enabling “the activation of all meaningful units of language” (Libben 2010; see also Dressler *et al.* 2003).

The offline and online experiments on adults and adolescents reported on here confirm that morphology may impact phonotactic processing inasmuch as homophonous sequences that either are monomorphemic or straddle a morphological boundary can be differently perceived, produced, manipulated. Additional testing with different experimental tasks and populations will provide more details on whether such parsing strategy is exclusive of early acquisition phases (or may also be effective in general processing), and on when and how the speakers’ sensitivity to the (mor)phonotactic status of clusters has facilitatory, inhibitory or no effects of phonological, morphological, lexical processing.

Addresses of the Authors

Chiara Celata, Laboratorio di Linguistica “G. Nencioni”, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56100 Pisa, Italy <c.celata@sns.it>

Katharina Korecky-Kröll, Department of Linguistics, University of Vienna, Porzellangasse 4/2/4, 1090 Vienna, Austria <katharina.korecky-kroell@univie.ac.at>

Irene Ricci, Laboratorio di Linguistics “G. Nencioni”, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56100 Pisa, Italy <i.ricci@sns.it>

Wolfgang U. Dressler, Department of Linguistics, University of Vienna, Porzellangasse 4/2/4, 1090 Vienna, Austria <wolfgang.dressler@univie.ac.at>

Acknowledgments

This research was funded in part by European Science Foundation “NetWordS” program, grant 4741-2012 to the first author (“Phonotactic effects on morphological structure”, <https://sites.google.com/site/morphotactics/>) and grant 6032-2013 to the second author (“New experimental

methods in research on morphonotactic processing”). The financial support for the realization of the experiments was provided by the Laboratorio di Linguistica of the Scuola Normale Superiore di Pisa. Thanks are due to all students of the “Psycholinguistisches Proseminar – Processing” course taught by the second author in summer term 2013 for their help with data collection, to Basilio Calderone for his participation in the discussion of both theoretical and experimental issues, to Sylvia Moosmüller for the recording of the test items, and to Giulia Bracco for her insightful comments on the behavioral experiments. Needless to say, any errors are the authors’ sole responsibility.

Bibliographical References

- Baayen Harald R., Piepenbrock Richard & Leon Gulikers 1995. *CELEX2*. Philadelphia: Linguistic Data Consortium.
- Bates Douglas, Maechler Martin, Bolker Ben & Steven Walker 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7.
- Bavin Edith (ed.) 2009. *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press.
- Calderone Basilio & Chiara Celata 2012. PHACTS about activation-based word similarity effects. In *Proceedings of the EACL 2012 Workshop on Computational Models of Language Acquisition and Loss*, The Association of Computational Linguistics. 12-19. <http://www.aclweb.org/anthology/W12-0900>
- Calderone Basilio, Celata Chiara, Korecky-Kröll Katharina & Wolfgang U. Dressler 2014. A computational approach to morphonotactics: evidence from German. *Language Sciences* 46. 59-70.
- Caramazza Alfonso, Laudanna Alessandro & Cristina Romani 1988. Lexical access and inflectional morphology. *Cognition* 28. 297-332.
- Celata Chiara, Calderone Basilio & Fabio Montermini 2011. Enriched sublexical representations to access morphological structure: A psycho-computational account. *TAL-Traitement Automatique du Langage* 52 (2). 49-71.
- Dressler Wolfgang U. & Katarzyna Dziubalska-Kołaczyk 2006. Proposing morphonotactics. *Italian Journal of Linguistics* 18(2), 102-144.
- Dressler Wolfgang U., Katarzyna Dziubalska-Kołaczyk & Lina Pestal 2010. Change and variation in morphonotactics. *Folia Linguistica Historica* 31. 51-68.
- Dressler Wolfgang U., Kilani-Schoch Marianne & Sabine Klampfer 2003. How does a child detect morphology? Evidence from production. In Baayen Harald R. & Robert Schreuder (eds), *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter. 391-425.
- Frauenfelder Uli H. & Ruth K. Kearns 1996. Sequence monitoring. *Language and Cognitive Processes* 11. 665-674.
- Freedman Skott E. & Jessica A. Barlow 2012 Using whole-word production measures to determine the influence of phonotactic probability and neighborhood density on bilingual speech production. *International*

- Journal of Bilingualism* 16. 369-387.
- Freiberger Eva Maria 2014. Der Erwerb von Konsonantenclustern im Deutschen: Eine Untersuchung longitudinaler Spontansprachkorpora. *Slowakische Zeitschrift für Germanistik* 6 (2). 5-23.
- Hay Jennifer 2003. *Causes and consequences of word structure*. London: Routledge.
- Kamandulyte Laura 2006. The acquisition of morphonotactics in Lithuanian. *Wiener Linguistische Gazette* 73. 88-96.
- Korecky-Kröll Katharina & Wolfgang U. Dressler forthcoming. (Mor)phonotactics in high vs. low SES children. Submitted to Dziubalska-Kołaczyk Katarzyna & Jarek Weckwerth (eds), *In memoriam of Rajendra Singh*. Poznań: Adam Mickiewicz University Press.
- Korecky-Kröll Katharina, Dressler U. Wolfgang, Freiberger Eva Maria, Reinisch Eva, Mörtz Karlheinz & Gary Libben 2014. Morphonotactic and phonotactic processing in German-speaking adults. *Language Sciences* 46. 48-58.
- Libben Gary 2010. Compounds, semantic transparency, and morphological transcendence. In S. Olson (ed.), *New impulses in word-formation (Linguistische Berichte Sonderheft 17)*. Hamburg: Buske. 317-330.
- Longtin Catherine-Marie, Segui Juan & Pierre A. Hallé (2003). Morphological priming without morphological relationship. *Language & Cognitive Processes* 18. 313-334.
- Marshall Chloe R. & Heather K.J. van der Lely 2006. A challenge to current models of past tense inflection: the impact of phonotactics. *Cognition* 100. 302-320.
- Mattys Sven L., Jusczyk Peter W., Luce Paul A. & James L. Morgan 1999. Word segmentation in infants: How phonotactics and prosody combine. *Cognitive Psychology* 38. 465-494.
- McQueen James M. 1998. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39. 21-46.
- Mehler Jacques, Dommergues Jean-Yves, Frauenfelder Uli & Juan Segui 1981. The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior* 20. 298-305.
- Plag Ingo, Homann Julia & Gero Kunter 2014. Homophony and morphology: the acoustics of word-final S in English. *Journal of Linguistics, FirstView Articles*. 1-36.
- Saffran Jenny R., Newport Elissa L. & Richard N. Aslin 1996. Word segmentation: the role of distributional cues. *Journal of Memory and Language* 35. 606-621.
- Song Yae Jung, Demuth Katherine, Evans Karen & Stefanie Shattuck-Hufnagel 2013. Durational cues to fricative codas in 2-year-olds' American English: Voicing and morphemic factors. *Journal of the Acoustical Society of America* 133 (5). 2931-2946
- Storkel Holly L. 2001. Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research* 44. 1321-1337.
- Taft Marcus & Sam Ardasinski 2006. Obligatory decomposition in reading prefixed words. *The Mental Lexicon* 1(2). 183-199.

- Taft Marcus & Kenneth I. Forster 1975. Lexical storage of polysyllabic words. *Journal of Verbal Learning and Verbal Behavior* 15. 607-620.
- Weissenborn Jürgen & Barbara Höhle (eds) 2000 *Approaches to Bootstrapping*. Amsterdam: Benjamins.
- Zydorowicz, Paulina 2010. Polish consonant clusters across morpheme boundaries: Polish morphonotactic inventory and its acquisition. *Poznan Studies in Contemporary Linguistics* 46. 565-588.

The interaction of prosody and phonotactics: Resyllabification in three varieties of German

Sylvia Moosmüller

Resyllabification is a syllable optimizing process, in which a coda consonant is assigned to the following syllable if the phonotactics of the language allows such reassignment. Since the process optimizes the structural well-formedness of the syllable, it is mostly observed in quantifying languages. The Middle Bavarian varieties of German feature some quantifying aspects as well. It was therefore tested whether, with respect to the word-final fricative /s/, resyllabification can be found in the Viennese dialect (VD) and in Standard Austrian German (SAG). The results were compared with results on Standard German (SGG), a language in which resyllabification is blocked by glottal stop insertion or glottalization in the onset. Results indicate that stress plays a crucial role: before a stressed word starting with a vowel, resyllabification is blocked by glottalization in all varieties, whereas before an unstressed word, absolute and relative durations rendered contradictory results. The results on absolute durations show a tendency towards resyllabification, whereas relative durations render significant differences between word-final /s/ in a context of resyllabification and word-initial /s/.

1. Introduction

Many researchers argue that CV structures constitute the optimal sequence of phonemes and constitute a language universal (e.g., Greenberg 1978; Vennemann 1988; Blevins 1995), since these structures are prevalent in all languages analysed to date. Yet, such a statement, though obviously true, inevitably entails the question of why so many language forms violate this optimal state. Morphology definitely plays a decisive role: Dziubaska-Kołaczyk (2014) shows that in Polish, consonant cluster size correlates with morphological complexity: “70 percent of the 2-consonant clusters are phonotactic while 70 percent of 4-consonant long clusters are morphologically complex” (2014: 13).

Thus, rich consonant clusters contain morphological information and make the (phonological) word more “visible” (Szczepaniak 2014), especially if they appear at the edges of the word. Therefore, they optimize the word by highlighting the word boundaries. On the other hand, in languages optimizing the syllable, word-boundaries are blurred and it is more difficult to decode morphological information. Therefore, an alternative approach to explain the predominance

of CV structures is provided by a typology based on prosody in which languages are classified according to whether they optimize the (phonological) word or the syllable. The approach of differentiating prototypical “word languages” and prototypical “syllable languages” has been elaborated by Auer (1991, 1993, 2001) as a continuation of the pure phonetically conceived concept of “stressed-timed” and “syllable-timed” languages (Pike 1945). Auer qualified the strong isochrony hypothesis advocated by Abercrombie (1967) by a) shifting the typological classification from the phonetic to the phonological level¹ and by b) proposing a continuum between the two prototypes and providing a list of mostly phonological parameters such as whether the sonority hierarchy is respected in consonant cluster sequences, whether vowels are reduced in non-accented syllables, or whether rules enhancing CV-structures or rich consonant sequences are observed (see Auer 1991: 303 for a complete list of processes).

Siebenhaar (2014) argues that it cannot be ultimately decided whether “segmental aspects result in a different rhythmic concept, or whether the rhythmic concept of a language entails the phonological structure” (2014: 332). However, whereas the first option has to rely on the admittedly intriguing but questionable concept of the well-known paradigm of ease of articulation and ease of perception in order to justify the preference for CV-structures, the second option provides a functional and more dynamic explanation for the continuous sliding of languages on a scale where either the word or the syllable is optimized. While this model still entails some shortcomings, such as too loosely defining the phonological word, and would profit from some refinements (Auer 2014), it still provides a typological explanation for why CV structures are common in all languages, whilst rich consonant clusters are predominantly found in “word” languages. Furthermore, it provides a good tool for handling mixed types such as the Alemannic (Siebenhaar 2014) or the Middle Bavarian (Bannert 1976, Moosmüller & Brandstätter 2014) varieties of German.

Typologically, Old High German (OHG) was a quantifying language, meaning that quantity opposition of both vowels and consonants was prevalent in all prosodic positions. On its way to New High German (NHG), unstressed syllables were gradually reduced, geminates were degeminated, long vowels became short vowels in unstressed positions, and short vowels were deleted (see Szczepaniak 2007 for a detailed overview of the processes which took place from OHG to NHG). Thus, German changed from a quantifying language to a word language, with rich consonant clusters and an uneven distribution of stressed and unstressed vowels (less vowel quality diver-

sity in unstressed positions, mostly reduced to schwa). However, these changes have also led to dialectal differences.

2. Description of the varieties

In the 19th century, the north German varieties spoken by the educated elites gained highest prestige and replaced the Upper Saxonian varieties which commanded this position at the times of Gottsched or Gellert. Thus, today, Standard German (SGG) is strongly associated with the variety spoken in the northern part of Germany, advocated in the media and ranked most prestigious in perception experiments (Herrgen 2015). However, some German scholars propose a broader definition of SGG, including regional standards. Therefore, it is predominantly the average educated speaker, independent of his or her regional origin, who is conceived as a speaker of SGG (Deppermann *et al.* 2013).

Among the varieties of German, SGG, both as defined in Herrgen (2015) and in Deppermann *et al.* (2013), with its rich consonant cluster inventory, its prevalence of schwa-vowels and the concomitant backgrounding of unstressed syllables, is located nearest to a prototypical word language.

The Middle Bavarian dialects, on the other hand, preserved some quantifying aspects. Though the number of consonants allowed in a cluster is at least as high or even higher than in SGG, e.g., [ˈqʃbr̥uːŋəl] *gesprungen* ‘spring, PRF’, [d͡ʒfriːʃ] *zu frisch* ‘too fresh’, [bʃviʃsɔ] *beschwipst* ‘tipsy’, the existence of full vowels in unstressed syllables and especially the specific timing of VC-sequences in stressed position still point to the preservation of old quantifying aspects in these varieties. Bannert (1976) and Ronneberger-Sibold (1999) assume an isochronous relationship between long vowel + lenis consonant sequences and short vowel + fortis consonant sequences, thus excluding further vowel + consonant combinations. However, Moosmüller (2007) and Moosmüller & Brandstätter (2014) demonstrate the existence of further combinations, especially long vowel + fortis plosive sequences. From their results, they favour a moraic approach in analysing stressed syllables in the Middle Bavarian dialects. Thus, the Middle Bavarian dialects show traits of both a word language and of a quantifying language (see also Siebenhaar 2014 for the Alemannic dialects).

The Viennese dialect (VD), which is analysed in the current contribution, belongs to the eastern Middle Bavarian dialects. As

an urban dialect, it holds the lowest prestige of all Austrian dialects (Moosmüller 1991). It is spoken by speakers of the lower social classes, mostly with a low educational background. Nevertheless, VD speakers learned a standard language in school and thus acquired passive, and, dependent on the speaker, a certain degree of active competence in SAG. Since many changes are currently observable in the speech of young speakers of the lower social classes (Glaunigner 2011), we will restrict VD speakers to speakers of the older generation, whose variety coincides with what is stereotypically considered as VD.

Among Austrian scholars, no uniform definition of Standard Austrian German (SAG) exists. As concerns the selections of speakers, some scholars rely on professional speakers (Muhr 2007; Wiesinger 2009), others on educated speakers not further specified (Bürkle 1995; Ehrlich 2009; Kleiner 2010). Some scholars restrict standard language use to specific text-types (predominantly reading of texts or official talks), others include spontaneous speech as well. These different approaches naturally result in different descriptions of SAG (see Moosmüller 2011, 2015).

Moosmüller (1991) based her definition of SAG on the results of perception and attitude experiments, conducted in the cities of Graz, Innsbruck, Salzburg, and Vienna. From these results it follows that above all, speakers from Salzburg and Vienna, to a lesser degree speakers from Graz, are accepted as speakers of SAG. Education turned out to be a decisive criterion: only speakers with a high educational background, whose parents also had a higher-level education, were accepted as SAG speakers. Most interestingly, Viennese speakers fulfilling these criteria were rated best. From these results it can be concluded that SAG as spoken in Vienna enjoys the highest prestige. These results were corroborated by more recent studies (Soukup 2009, Goldgruber 2011).

A historical view on the development of SAG reveals that the Middle Bavarian variety spoken by the nobility and educated elite of the 18th century held a very low prestige (Wiesinger 1985). Thus, speakers of these social classes geared towards the most prestigious variety spoken in the respective German countries of the time (see Moosmüller 2015 for an overview). Consequently, SAG holds an in-between position, showing both traits of its Middle Bavarian basis, especially as concerns phonetics and phonology, and of SGG, especially as concerns morphology, syntax, and the lexicon.² The congruities of SGG and SAG on higher grammatical levels severely challenge the quantity approach advocated by Bannert (1976) and Ronneberger-

Sibold (1999), since an abundant number of lexical items violate the phonotactic restrictions put forward in their treatises. An integration of SGG lexical items into the Middle Bavarian timing patterns is more easily achieved in the mora-based approach proposed by Moosmüller & Brandstätter (2014), which challenges the traditional view that only long vowel + lenis consonants or short vowel + fortis consonants are allowed in the Middle Bavarian varieties and prove the existence of a third category, namely of long vowel + fortis consonant sequences.

From the results described so far, it can be concluded that in SGG, the word is optimized, whereas in VD and SAG, both word optimizing and syllable optimizing tendencies can be observed.

3. Resyllabification

Resyllabification constitutes a further parameter which optimizes the syllable and consequently blurs word boundaries. As such, resyllabification is most extensively described for languages which are generally thought of to belong to a syllable-optimizing type such as Spanish (Torreira & Ernestus 2012, Strycharczuk *et al.* 2014), Catalan (Caro Reina 2014), French (Gaskell *et al.* 2002, Fougeron 2007), Italian (Loporcaro & Bertinetto 2005), and Alemannic (Siebenhaar 2014).

In a process of resyllabification, coda consonants are assigned to the following syllable if the phonotactics of the language allow such reassignment. A word-final consonant followed by a word-initial vowel would constitute a very typical context for resyllabification, termed *enchaînement* in French. In the case of resyllabification, the coda consonant is assumed to adopt the traits of an onset consonant and thus impede word recognition. Fougeron *et al.* (2003) looked for acoustic cues of resyllabification and analysed three conditions: a) condition of enchaînement ($V_1.C\#V_2$), b) word boundary condition ($V_1\#CV_2$), and c) syllable boundary condition ($V_1.CV_2$). However, with speaker-specific differences, they found significantly longer durations of the vowel preceding the consonant in the enchaînement condition and in the vowel in the word boundary condition as compared to the vowel in the syllable boundary condition. In a condition of resyllabification or enchaînement,³ the vowel preceding the consonant turns into a word-final vowel and is subject to word-final lengthening. As concerns the consonants, Fougeron *et al.* (2003) obtained different results dependent on consonant type. While stops in the enchaînement conditions

were realized shorter than in both the word boundary and the syllable boundary conditions, no such differences occurred with respect to fricatives and liquids. These differences are also reflected in the duration of V_2 : vowels following stops are shorter in the enchaînement condition, whereas vowels following fricatives show no differences regarding the three boundary conditions. From their results, Fougeron *et al.* (2003) conclude that enchaînement in French is not simply a resyllabification of the word-final coda to the onset of the following syllable.

Similar results have been obtained for French by Dumay *et al.* (1999), Spinelli *et al.* (2002), Gaskell *et al.* (2002), and Fougeron (2007), and for Spanish varieties by Lipski (1999) and Strycharczuk & Kohlberger (2014).

For Alemannic, Siebenhaar (2014) could show that word-initial fricatives are significantly longer than word-medial and word-final fricatives. In a condition of resyllabification, i.e., a word-final fricative preceding a word-initial nasal, liquid, or vowel, the word-final fricative is lengthened and reaches the mean duration of the word-initial fricative (2014: 339). Thus, in Alemannic, full resyllabification is observed, at least as far as durational parameters are concerned.

In SGG, glottal stop insertion before word-initial vowels usually blocks resyllabification. However, Bergmann (2014) showed that boundary weakening, i.e., glottalization,⁴ sometimes takes place in unaccented positions and in high frequency words. In Alemannic, glottal stop insertion does not exist, and is therefore open for resyllabification. In the Middle Bavarian varieties, glottal stop insertion is optional, and mostly replaced by glottalization. In addition, word boundary marking can be further weakened and be signalled by a drop in F_0 or a drop in amplitude (Moosmüller 1999). Therefore, a gradual pacing from glottal stop to a loss of marking the word boundary can be observed, the latter case constituting a condition for resyllabification.

In the current contribution, the fricative /s/ is analysed with respect to resyllabification. For /s/, however, variety-specific outputs for resyllabification have to be assumed for the three varieties under consideration. While SGG features voiced [z] word-initially, but voiceless [s] word-finally, in a condition of resyllabification, the resyllabified fricative should surface as voiced [z], accompanied by a shortening in duration in order to equal the durations of word-initial [z]. In the Middle Bavarian varieties, on the other hand, both word-final and word-initial /s/ is voiceless. Therefore, in a condition of resyllabification, the resyllabified fricative should, in accordance with Siebenhaar (2014), show the same durational traits as a word-initial fricative.

Therefore, it is hypothesized that

1. for all varieties, glottal stop insertion or glottalization blocks resyllabification before stressed syllables,
2. glottal stop insertion or glottalization blocks resyllabification in SGG,
3. for the Middle-Bavarian varieties, resyllabification takes place before unstressed syllables,
4. resyllabification takes place more readily in VD than in SAG.

4. Method

Eight speakers (4 females and 4 males, age range 45-60 years) of the Viennese dialect, 16 speakers of SAG (4 females and 4 males, age range 45-60 years, and 4 females and 4 males, age range 18-25 years), and 4 speakers of SGG (3 females and 1 male, age range 25-59 years)⁵ were asked to read a list of 72 sentences twice.

The Austrian subjects were raised in Vienna with at least one parent raised in Vienna as well. The dialect speakers have finished secondary school, but not all of them have finished an apprenticeship. The SAG speakers of the older group hold an academic degree. The younger speakers have finished grammar school with 'Matura' (Austrian high school diploma/school leaving certificate) and were pursuing a university education at the time of recording. At least one parent of each SAG speaker held an academic degree as well.

The German subjects hold an academic degree and were chosen according to the criteria proposed in Deppermann *et al.* (2013), which defines a standard speaker as an averagely educated speaker.

From the list of read sentences, the following duration measurements of the fricative /s/ were performed:⁶

- the word-initial /s/, in a stressed position of a content word (n = 168): e.g., *Saft* 'juice', *Siebe* 'sieves',
- the word-initial /s/, in an unstressed syllable of a content word (n = 168): e.g., *Salat* 'salad', *Sekretär* 'secretary'
- the word-initial /s/, in an unstressed position of a pronoun (n = 168): e.g., *wie sich der Kater* 'as the cat itself', *können Sie* 'you can', *ergriff sie* 'she was grasped',
- the word-final /s/ of a stressed word (n = 56): e.g., *mit Kies bestreut* 'strewn with gravel',
- the word-final /s/, followed by a stressed vowel (V₁s#V₂) (n = 56): e.g., *köstliches Essen* 'delicious food' = resyllabification Condition 1,
- the word-final /s/, followed by an obstruent (V₁s#OV₂) (n = 56):

e.g., *köstliches Bier* ‘delicious beer’,

- the word-final /s/ of an unstressed word, followed by an unstressed word (n = 56): e.g., *aus der* ‘from the’,
- the word-final /s/, followed by an unstressed vowel ($V_1s\#V_2$) (n = 56): e.g., *zünd nicht das Haus an* ‘don’t torch the house’ = resyllabification Condition 2,
- V_1 in the conditions ($V_1s\#V_2$), ($V_1s\#OV_2$), and ($V_1s\#V_2$).
 Durations of the fricative were normalized for word duration.

5. Results

5.1. /s/ in word-initial position

For word-initial /s/, three positions were tested: word-initial /s/ for stressed content words, word-initial /s/ for content words in which the first syllable is unstressed, and word-initial /s/ of pronouns in an unstressed position. Figure 1 reveals that, unsurprisingly, the absolute duration of /s/ in unstressed pronouns is shortest, whereas, on the other hand, surprisingly, the relative duration of word-initial /s/ in unstressed pronouns is higher than in word-initial /s/ in a stressed position or in an unstressed position of a content word, thus taking up more than half of the total word duration.

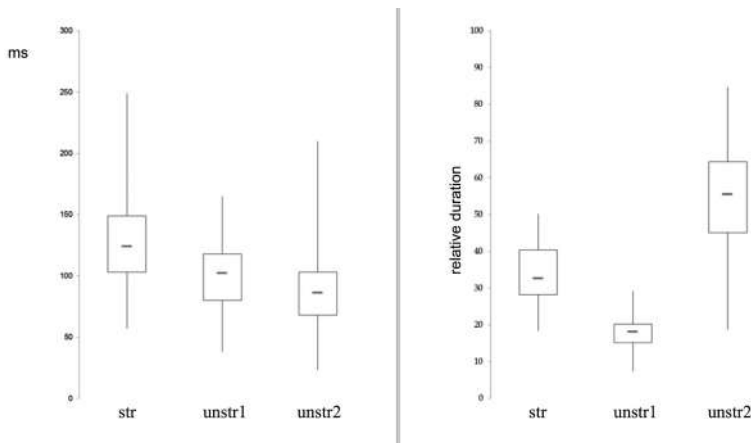


Figure 1. Boxplots of absolute duration (left) and relative duration (right) of word-initial /s/ in a stressed position (=str), in an unstressed position of a content word (=unstr1), and in unstressed pronouns (=unstr2) for all groups of speakers.

However, the results of a two-way ANOVA (speaker group x type of /s/) not only disclose a highly significant main effect for the three types of /s/ [$F(2,483) = 766.49$, $p < .0001$], but also a highly significant main effect for the speaker group [$F(6,483) = 7.24$, $p < .0001$]. Figure 2 presents the results broken down by speaker group. Because of space constraints and since the results on the word-initial /s/ in unstressed pronouns are the most relevant in the context of resyllabification Condition 2, only these results will be presented in Figure 2.

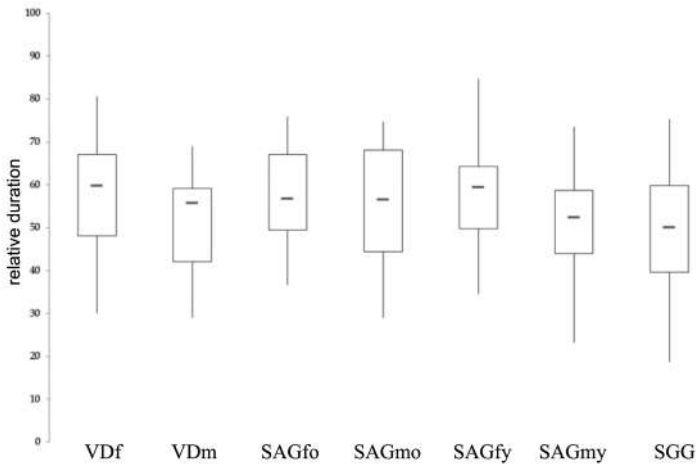


Figure 2. Boxplots of the relative duration of word-initial /s/ in unstressed pronouns, broken down by groups of speakers. VD = Viennese Dialect, SAG = Standard Austrian German, SGG = Standard German German, f = female, m = male, o = older generation, y = younger generation.

With respect to the relative durations of word-initial /s/ in unstressed pronouns, young male SAG speakers behave similarly to SGG speakers. This might be due to the fact that in two of the three items tested, namely in *wie sich der Kater* ‘as the cat itself’ and in *können Sie* ‘you can’, the /s/ is preceded by a vowel/voiced sonorant. Although the phoneme inventories of SAG and VD feature no voiced /z/ in word-initial positions, voiced contexts might cause a voicing of the fricative. Figure 3 shows the percentage of voicing in these two conditions, broken down by speaker group.

As expected, /z/ is voiced in SGG.⁷ In SAG, voicing of /s/ occurs predominantly in the production of male SAG speakers. However,

the percentage of voiced [z] in young male SAG speakers is not high enough to justify the assumption that the comparable results in the relative durations of young male SAG speakers and SGG speakers is due to voicing.

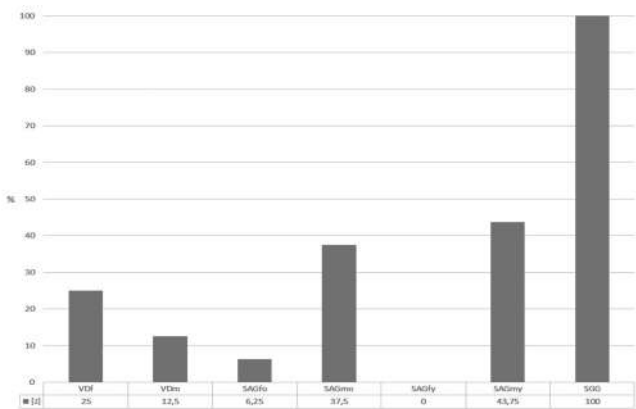


Figure 3. Percentage of voicing of word-initial /s/ in the contexts *wie sich der Kater* ‘the cat itself’ and *können Sie* ‘you can’. VD = Viennese Dialect, SAG = Standard Austrian German, SGG = Standard German German, f = female, m = male, o = older generation, y = younger generation.

5.2. /s/ in word-final position ($V_1s\#OV_2$)

Three positions were tested for word-final /s/: the word-final /s/ of a stressed word followed by a word-initial consonant of an unstressed syllable, the word-final /s/ of an unstressed word followed by a word-initial consonant of a stressed word, and the word-final /s/ of an unstressed word followed by a word-initial consonant of an unstressed word. A comparison of absolute and relative durations yielded differences with respect to the stressed vs. unstressed condition followed by a stressed word (see Figure 4). In these two conditions, absolute durations disclose no significant differences [$t(119) = 0.13$, $p = .45$], whereas relative durations yielded a highly significant difference [$t(85) = 24.16$, $p < .0001$]. This result is interesting, since it indicates that relative durations are sensitive to the number of syllables (one syllable in *Kies* ‘gravel’, three syllables in *köstliches* ‘delicious’), whereas absolute durations are not influenced by the composition of the word.

As concerns unstressed words followed by an unstressed word, results resemble those for absolute and relative durations in unstressed word-initial position: absolute durations are shortest, while relative durations are highest. However, the results of a one-tailed t-test revealed a highly significant difference [$t(129) = 12.24$, $p < 0.0001$] between the relative duration of a word-initial unstressed /s/ in pronouns and a word-final unstressed /s/ followed by unstressed words. Relative duration is lower in a word-final unstressed position followed by an unstressed word than in the word-initial position of unstressed pronouns (see right figure of Figure 1 and right figure of Figure 4 for comparison).

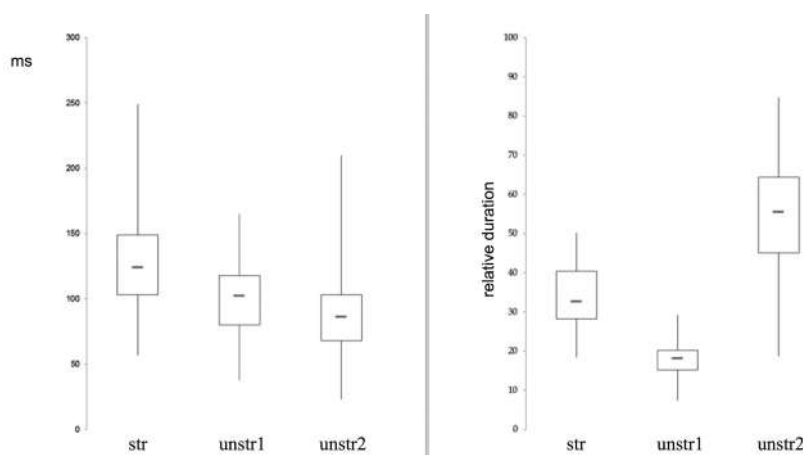


Figure 4. Boxplots of absolute duration (left) and relative duration (right) of a word-final /s/ in a stressed position (=str), in an unstressed position followed by a stressed word (=unstr1), and in an unstressed word followed by an unstressed word (= unstr2) for all groups of speakers.

A two-way ANOVA (speaker group x type of /s/) revealed a highly significant main effect for the three types of /s/ [$F(2, 147) = 210.13$, $p < .0001$] and a highly significant main effect for the speaker group [$F(6, 147) = 4.19$, $p < .001$]. Figure 5 presents the results on relative durations of word-final stressed /s/, broken down by speakers, since this condition is the most relevant for a comparison with resyllabification Condition 2.

As becomes apparent from Figure 5, no significant differences exist between the SGG speakers and the young SAG speakers [$t(22) = 0.29$, $p = .39$]. The same holds for absolute durations [$t(22) = 0.40$, $p = .35$].

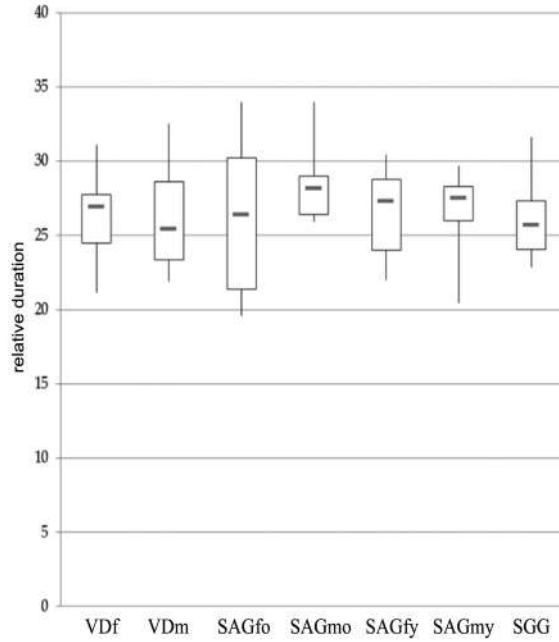


Figure 5. Boxplots of the relative duration of word-final /s/ in stressed words, broken down by groups of speakers. VD = Viennese Dialect, SAG = Standard Austrian German, SGG = Standard German German, f = female, m = male, o = older generation, y = younger generation.

5.3. /s/ in resyllabification Condition 1: ($V_1s\#V_2$)

In this condition, resyllabification is blocked by glottal stop insertion or by glottalization of the stressed vowel. Bergmann (2014) differentiates between glottal stop insertion and glottalization, since the latter constitutes a weakening of glottal stop insertion. However, in the current contribution, both are treated as equal, since both block resyllabification. Since glottal stops are rare in the Austrian varieties, glottalization occurs in the majority of cases. Therefore, the term “glottalization” will be used for both glottal stop insertion and glottalization.

In the $V_1s\#V_2$ condition, glottalization is suppressed in four items only, namely by one VD speaker, by one female SAG speaker of the older generation (twice) and by one male SAG speaker of the older generation.

Since the condition for resyllabification is met only in the four cases mentioned, the durations of /s/ in the condition ($V_1s\#V_2$) and ($V_1s\#OV_2$), *köstliches Essen* and *köstliches Bier*, respectively, should

remain the same. With respect to absolute duration, this is in fact the case. A two-way ANOVA yielded no significant result for the type of /s/; [$F(1, 98) = 3.51, p = .06$]. However, with respect to relative duration, a highly significant difference occurred [$F(1, 98) = 10.76, p < .001$], as shown in Figure 6.

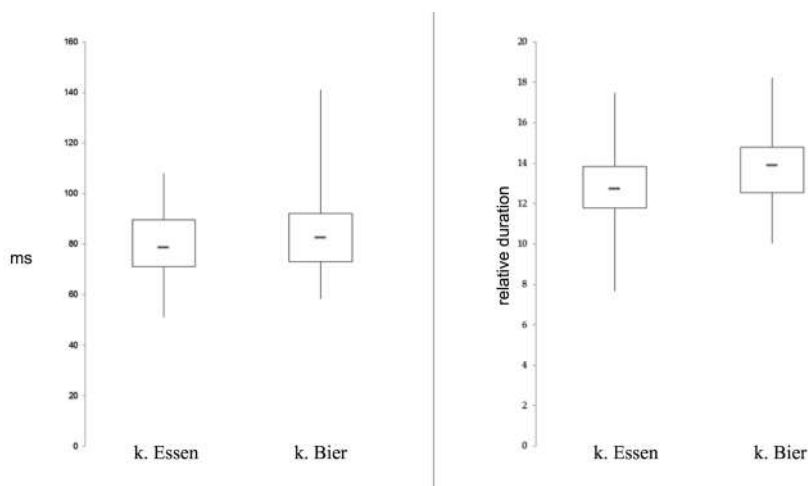


Figure 6. Boxplots of absolute duration (left) and relative duration (right) of the word-final /s/ in *köstliches Essen* (= k. Essen) and *köstliches Bier* (= k. Bier), for all groups of speakers.

Both the results on relative and absolute duration yielded a highly significant main effect for the speaker group, but only the results on absolute duration rendered a significant interaction for the type of /s/ and the speaker group [$F(6,98) = 2.4, p < .04$]. Pronounced differences between the two types of /s/ are obvious in SAG speakers of the older generation, whereas no such differences are observable in the SGG speakers and the SAG speakers of the younger generation, and much smaller differences are visible in the VD speakers.

5.4. /s/ in resyllabification Condition 2: ($V_1s\#V_2$)

Things are quite different in resyllabification Condition 2 as compared to resyllabification Condition 1. In the case of a word-final /s/ followed by an unstressed word-initial vowel, glottalization is suppressed

in 40% of all observed $V_1s\#V_2$ sequences. Again, some group-specific differences appear (see Figure 7), especially within the group of VD speakers. While all male VD speakers suppress glottalization in this condition, this is not the case among the female VD speakers. Most interestingly, one SGG speaker also suppresses glottalization in this condition.

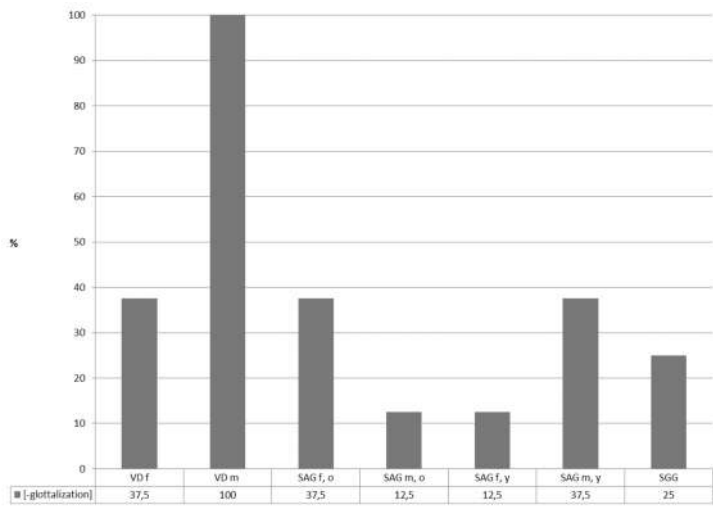


Figure 7. Suppression of glottalization (in %) in the condition $V_1s\#V_2$. VD = Viennese Dialect, SAG = Standard Austrian German, SGG = Standard German German, f = female, m = male, o = older generation, y = younger generation.

Consequently, in the case where glottalization takes place, /s/ is supposed to stay in the word-final position and should thus remain on par in duration with the word-final /s/ of a stressed word. In the case of a suppression of the process, a word-final /s/ is supposed to move to the onset of the next syllable and should thus be on par in duration with a word-initial /s/ in unstressed pronouns. For SGG speakers, resyllabified /s/ should surface as voiced [z]. Since this is not the case, it can be assumed that no resyllabification takes place in SGG. Consequently, SGG speakers are excluded from the following calculations.

However, t-tests rendered different results according to whether absolute or relative durations were tested. In the condition where

resyllabification is blocked by glottalization, absolute durations of /s/ in *Haus* ‘house’ and /s/ in *Kies* ‘gravel’ are the same [$t(78) = 0.25$, $p = .803$]. However, highly significant differences occur with respect to relative duration [$t(78) = -8.48$, $p < .0001$]; the relative duration of word-final /s/ in *Haus* being lower than of word-final /s/ in *Kies*. Figure 8 compares the results of absolute and relative durations for the [+glottalization] condition.

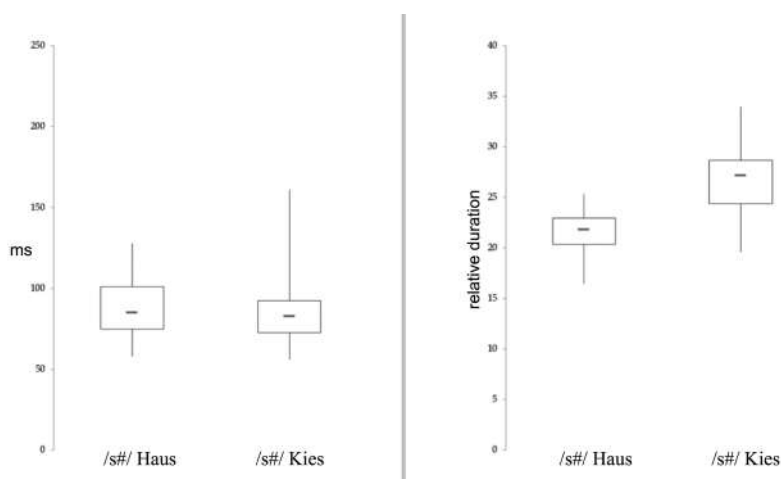


Figure 8. Boxplots of absolute duration (left) and relative duration (right) of word-final /s/ of Haus in the [+glottalization] condition (= /s#/ Haus) and of word-final /s/ of Kies (= /s#/ Kies), for all groups of Austrian speakers.

In the other condition, where resyllabification is not blocked by glottalization, absolute and relative durations of the word-final /s/ of *Haus* and of the word-initial /s/ of unstressed pronouns show statistically significant differences. Absolute durations are longer in word-initial unstressed pronouns than in the word-final /s/ of *Haus* [$t(156) = -1.65$, $p = .05$]. Highly significant differences occur for relative durations [$t(51) = -25.01$, $p < .0001$]; relative durations being higher in word-initial unstressed pronouns. Figure 9 compares the results of absolute and relative durations for the [-glottalization] condition.

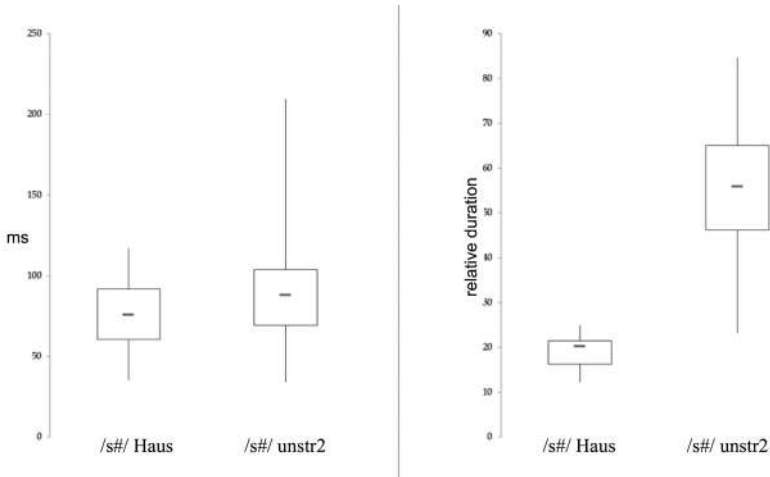


Figure 9. Boxplots of absolute duration (left) and relative duration (right) of word-final /s/ of Haus in the [-glottalization] condition (= /s#/ Haus) and of word-initial /s/ of unstressed pronouns (= /s#/ unstr2), for all groups of Austrian speakers.

A comparison of the [+glottalization] and the [-glottalization] condition of the word-final /s/ of *Haus* revealed minor differences in absolute durations [$t(18) = 1.48$, $p = 0.08$]; /s/ tends to be shorter in the [-glottalization] condition. Differences in relative durations are significant [$t(17) = 2.02$, $p = .03$], with relative duration being lower in the [-glottalization] condition. The diphthong of *Haus* tends to be longer in the [-glottalization] condition [$t(46) = -1.41$, $p = .08$], however, with respect to relative durations, highly significant differences occur [$t(46) = -3.31$, $p < .001$]. Relative durations are higher in the [-glottalization] condition as compared to the [+glottalization] condition.

6. Discussion

In this contribution, it was tested whether resyllabification takes place across a word-boundary when the following word starts with a vowel, or whether resyllabification is blocked by glottal stop insertion or glottalization. The realization of resyllabification would further undermine quantifying aspects of VD and SAG, since resyllabification is a syllable optimizing process. It was hypothesized that resyllabification would more readily take place in VD and SAG as compared to

SGG. This general hypothesis was confirmed; in SGG, resyllabification is blocked by glottalization. In the unstressed context of $V_1s\#V_2$, one SGG speaker suppresses glottalization. However, since the word-final /s/ is not pronounced as a voiced [z] as demanded by the phonotactics of SGG in a word-initial position, it is assumed that /s/ is not resyllabified.

In the condition ($V_1s\#V_2$), i.e., in the condition of a word-final /s/ with a subsequent word-initial stressed vowel, resyllabification is blocked by glottalization in all varieties under consideration. In the second condition ($V_1s\#V_2$), i.e., in the condition of a word-final /s/ with a subsequent word-initial unstressed vowel, glottalization is suppressed in 40% of all cases.

When glottalization is suppressed, i.e., in the [-glottalization] condition, only minor differences occur with respect to absolute durations of the word-final /s/ of *Haus* and the word-initial /s/ of unstressed pronouns. In addition, the diphthong of *Haus* shows a tendency to be longer in the [-glottalization] condition than in the [+glottalization] condition. Therefore, if absolute durations are considered, a tendency to resyllabification can be observed.

These results are not corroborated when relative durations are considered. The relative duration of the word-final /s/ of *Haus* is significantly lower than the relative duration of the word-initial /s/ of unstressed pronouns. However, the relative duration of the diphthong in the [-glottalization] condition is significantly higher than in the [+glottalization] condition. Considering relative durations, only the results of the diphthong duration indicate a resyllabification. The relative duration of word-final /s/ is both lower than the duration of the word-initial unstressed pronouns and the word-final /s/ in the [+glottalization] condition, establishing the following hierarchy: word-initial /s/ of unstressed pronouns > word-final /s/ in the [+glottalization] condition > word-final /s/ in the [-glottalization] condition.

Differences with respect to absolute and relative durations were also observed in resyllabification Condition 1: considering the absolute duration of the word-final /s/ of *köstliches*, no differences occurred with respect to the following context (consonant vs. vowel), as expected, whereas a comparison of the relative durations rendered highly significant differences.

These results address the question of whether absolute or relative durations have to be considered in clarifying whether resyllabification occurs in the Austrian varieties analysed. However, this question can only be answered by conducting perception experiments, after which, an extension of the analysis to spontaneous speech is

meaningful and will further reveal whether, with respect to resyllabification, the Middle Bavarian varieties of German show quantifying or word-optimizing traits.

Results also revealed similarities between SGG speakers and young SAG speakers. Currently, immigrants from Germany constitute the largest immigrant group in Austria, especially in Vienna. Since SGG is the most prestigious standard variety of the German speaking countries, a contact-induced sound change seems natural. Therefore, further investigations to identify a sound change in progress and find out if this sound change is influenced by SGG will be dealt with in future studies.

Address of the Author

Sylvia Moosmüller, Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, A-1040 Wien, Austria <sylvia.moosmuller@oeaw.ac.at>

Acknowledgements

The current investigation was performed within the project I 1394-G23 “Human Behavior and Machine Simulation in the Processing of (Mor) Phonotactics”, funded by the FWF. I am especially grateful to two reviewers for their helpful comments on a previous version of this contribution and to Jonathan White for the correction of the English.

Notes

¹ The correlation of rhythm and language structure goes back to Donegan & Stampe (1983).

² Generally, works discussing SAG lexical entries mostly discuss the differences between SAG and SGG lexicon; hardly ever do they focus on differences between SAG and the dialects.

³ Except for the discussion and conclusions part, Fougeron *et al.* (2003) use the terms “enchaînement” and “resyllabification” synonymously.

⁴ The boundary is signalled by creaky voice in the vowel. Acoustically, it is sometimes difficult to detect the exact point in time of the boundary.

⁵ The results from Bergman (2014) proved that resyllabification is blocked in SGG. Therefore, the group of SGG speakers only served as a control group and was thus not balanced for gender or age.

⁶ The software STx has been used, see <http://www.kfs.oeaw.ac.at>

⁷ In the third context, namely *ergriff sie* ‘she was grasped’, /z/ is voiceless in SGG. Contrary to SAG and VD, a process of devoicing has to be assumed in SGG.

Bibliographical References

- Abercrombie David 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- Auer Peter 1991. "Stress-Timing" vs. "Syllable-Timing" from a Typological Point of view. In Palek Bohumil & Přemysl Janota (eds.), *Proceedings of the Conference Linguistics and Phonetics: prospects and applications*. Prague: Charles University Press. 292-305.
- Auer Peter 1993. Is a rhythm-based typology possible? A study of the role of prosody in phonological typology. *KonTRI Working Papers 21*. (http://paul.igl.uni-freiburg.de/auer/userfiles/downloads/Phonotypo_Kontri1.pdf).
- Auer Peter 2001. Silben- und akzentzählende Sprachen. In Haspelmath Martin, König Ekkehard, Oesterreicher Wulf & Wolfgang Raible (eds.), *Language Typology and Language Universals*. Berlin/New York: Walter de Gruyter. 1391-1399.
- Bannert Robert 1976. *Mittelbairische Phonologie auf akustischer und perzeptorischer Grundlage*. Lund: Gleerup.
- Bergmann Pia. 2014. Reduction and deletion of glottal stops and geminates at phonological word boundaries in German compounds: Effects of word frequency and accentuation. In Caro Reina Javier & Renata Szczepaniak (eds.), *Syllable and Word languages*. Berlin/New York: Walter de Gruyter. 251-278.
- Blevins Juliette 1995. The syllable in phonological theory. In Goldsmith John (ed.), *The Handbook of Phonological Theory*. Cambridge: Wiley-Blackwell. 206-244.
- Bürkle Michael 1995. *Zur Aussprache des österreichischen Standarddeutschen*. Frankfurt/Main: Lang.
- Caro Reina Javier 2014. Central Catalan in the framework of the typology of syllable and word languages. In Caro Reina Javier & Renata Szczepaniak (eds.), *Syllable and Word languages*. Berlin/New York: Walter de Gruyter. 349-390.
- Deppermann Arnulf, Kleiner Stefan & Ralf Knöbl 2013. 'Standard usage': Towards a realistic conception of spoken standard German. In Auer Peter, Caro Reina Javier & Göz Kaufmann (eds.), *Language Variation – European Perspectives IV. Selected Papers from the Sixth International Conference on Language Variation in Europe (ICLaVE)*. Amsterdam: John Benjamins. 83-116.
- Donegan Patricia & David Stampe 1983. Rhythm and the holistic organization of language structure. *Chicago Linguistic Society* 19. 337-353.
- Dumay Nicolas, Content Alain & Ulrich Frauenfelder 1999. Acoustic-Phonetic Cues To Word Boundary Location: Evidence From Word Spotting. In Ohala John, Hasegawa Yoko, Ohala Manjari, Granville Daniel & Ashlee C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences*. San Francisco. 281-284.
- Dziubaska-Kołaczyk Katarzyna 2014. Explaining phonotactics using NAD. *Language Sciences* 46. 6-17.
- Ehrlich Karoline 2009. *Die Aussprache des Österreichischen Standarddeutsch – umfassende Sprech- und Sprachstandserhebung der österreichischen*

- Orthoëpie*. PhD thesis, University of Vienna.
- Fougeron Cécile 2007. Word boundaries and contrast neutralization in the case of enchaînement in French. In Cole Jennifer & José Ignacio Hualde (eds.), *Laboratory Phonology 9*. Berlin, New York: Walter de Gruyter. 609-642.
- Fougeron Cécile, Bagou Odile, Content Alain, Stefanuto Muriel & Ulrich Frauenfelder 2003. Looking for acoustic cues of resyllabification in French. In Solé Maria-Josep, Recasens Daniel & Joaquim Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS 2003)*. Barcelona. 2257-2260.
- Gaskell M. Gareth, Spinelli Elsa & Fanny Meunier 2002. Perception of resyllabification in French. *Memory & Cognition* 30(5). 798-810.
- Glauninger Manfred 2011. "Wiener(isch) zwischen Wunschbild und Wirklichkeit. Dialekt im Wienerlied und im Wiener Sprachalltag". *Jahrbuch des Österreichischen Volksliedwerkes* 60. 142-153.
- Goldgruber Barbara 2011. *Einstellungen zu Dialekt und Standardsprache in Österreich. Eine empirische Untersuchung in Graz und Wien*. Master thesis, University of Vienna.
- Greenberg Joseph 1978. Some generalizations concerning initial and final consonant clusters. In Id. (ed.), *Universals of Human Language. Volume 2. Phonology*. Stanford: Stanford University Press. 243-280.
- Herrgen Joachim 2015. Entnationalisierung des Standards. Eine perzeptionslinguistische Untersuchung zur deutschen Standardsprache in Deutschland, Österreich und der Schweiz. In Glauninger Manfred & Alexandra Lenz (eds.), *Standarddeutsch in Österreich – Theoretische und empirische Ansätze*. Wien: Vandenhoeck & Ruprecht. 137-162.
- Kleiner Stefan 2010. Zur Aussprache von nebentonigem -ig im deutschen Gebrauchsstandard. *Zeitschrift für Dialektologie und Linguistik* LXXVII, 3. 259-303.
- Lipski John M. 1999. The Many Faces of Spanish /s/-Weakening: (Re)alignment and Ambisyllabicity. In Gutiérrez-Rexach Javier & Fernando Martínez-Gil (eds.), *Advances in Hispanic Linguistics*. Somerville, MA: Cascadia Press. 198-213.
- Loporcaro Michele & Pier Marco Bertinetto 2005. The sound pattern of Standard Italian, as compared with varieties spoken in Florence. *Journal of the International Phonetic Association* 35(2). 131-151.
- Moosmüller Sylvia 1991. *Hochsprache und Dialekt in Österreich. Soziophonologische Untersuchungen zu ihrer Abgrenzung in Wien, Graz, Salzburg und Innsbruck*. Wien/Köln/Weimar: Böhlau.
- Moosmüller Sylvia 1999. The modulation of formant amplitude during assimilation processes of homorganic and quasi-homorganic V1V2-sequences. In Ohala John, Hasegawa Yoko, Ohala Manjari, Granville Daniel & Ashlee C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences*. San Francisco. 1329-1332.
- Moosmüller Sylvia 2007. On some timing aspects in the Viennese dialect. *The Phonetician* 95. 19-27.
- Moosmüller Sylvia 2011. Aussprachevarianten im Österreichischen Standarddeutsch. In Bose Ines & Baldur Neuber (eds.), *Interpersonelle Kommunikation: Analyse und Optimierung*. Frankfurt/Main: Lang. 193-199.

- Moosmüller Sylvia 2015. Methodisches zur Bestimmung der Standardaussprache in Österreich. In Glauninger Manfred & Alexandra Lenz (eds.), *Standarddeutsch in Österreich – Theoretische und empirische Ansätze*. Wien: Vandenhoeck & Ruprecht. 163-182.
- Moosmüller Sylvia & Julia Brandstätter. 2014. Phonotactic information in the temporal organization of Standard Austrian German and the Viennese dialect. *Language Sciences* 46. 84-95.
- Muhr Rudolf. 2007. *Österreichisches Aussprachewörterbuch. Österreichische Aussprachedatenbank*. Frankfurt/Main: Peter Lang.
- Pike Kenneth L. 1945. *The Intonation of American English*. Ann Arbor: University of Michigan Press.
- Ronneberger-Sibold Elke 1999. Ambisyllabic consonants in German: Evidence from dialectal pronunciation of lexical creations. In Rennison John & Klaus Kühnhammer (eds.), *Phonologica 1996*. The Hague: Theseus. 247-271.
- Siebenhaar Beat 2014. Phonological and phonetic considerations for a Classification of Swiss German dialects as a word language or a syllable language. In Caro Reina Javier & Renata Szczepaniak (eds.), *Syllable and Word languages*. Berlin/New York: Walter de Gruyter. 327-345.
- Soukup, Barbara 2009. *Dialect use as interaction strategy. A sociolinguistic study of contextualization, speech perception, and language attitudes in Austria*. Wien: Braumüller.
- Spinelli Elsa, McQueen James M. & Anne Cutler 2003. Processing resyllabified words in French. *Journal of Memory and Language* 48. 233-254.
- Strycharczuk Patrycja, van 't Veer Marijn, Bruil Martine & Linke Kathrin 2014. Phonetic evidence on phonology-morphosyntax interactions: Sibilant voicing in Quito Spanish. *Journal of Linguistics* 50(2). 403-452.
- Strycharczuk Patrycja & Martin Kohlberger 2014. Resyllabification reconsidered. On the phonetic properties of derived onsets in Spanish. Presentation given at the 11th Old World Conference in Phonology (OCP11), January 22-25, 2014, Leiden University Centre for Linguistics.
- Szczepaniak Renata 2007. *Der phonologisch-typologische Wandel des Deutschen von einer Silben- zu einer Wortsprache*. Berlin: Mouton de Gruyter.
- Szczepaniak Renata 2014. Vowel and consonant epenthesis in the history of German from the typological perspective of syllable and word languages. In Caro Reina Javier & Renata Szczepaniak (eds.), *Syllable and Word languages*. Berlin/New York: Walter de Gruyter. 160-180.
- Torreira Francisco & Mirjam Ernestus 2012. Weakening of intervocalic /s/ in the Nijmegen Corpus of Casual Spanish. *Phonetica* 69. 124-148.
- Vennemann Theo 1988. *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin: Mouton de Gruyter.
- Wiesinger Peter 1985. Die Diagliederung des Deutschen seit dem 17. Jahrhundert. In Besch Werner, Betten Anne, Reichmann Oskar & Stefan Sonderegger (eds.) *Sprachgeschichte*. Bd. 2. Berlin: de Gruyter. 1633-1650.
- Wiesinger Peter 2009. Die Standardaussprache in Österreich. In Krech Eva-Maria, Stock Eberhard, Hirschfeld Ursula & Lutz Christian Anders (eds.), *Deutsches Aussprachewörterbuch*. Berlin: de Gruyter. 229-258.

Why do syllable onsets attract consonant(s)?

Nathalie Vallée, Thi Thuy Hien Tran, Paolo Mairano, Silvain Gerber & Solange Rossato

The present study is carried out in the context of a project whose primary aim is to define the physical components that could help explain what makes a syllable. We examine whether phonetic factors can account for observed universal trends in syllable structure, particularly those that provide an explanation for the Maximum Onset Principle (MOP). MOP forms the basis of syllabification theories allocating consonantal segments to syllable-initial position rather than to syllable-final position. Two theoretical frameworks can provide a phonetic basis for MOP, namely the Frame/Content Theory and Articulatory Phonology, but they do not account for it thoroughly. Consequently, we propose to reconsider the MOP in the light of new articulatory data focusing on basic jaw movements, and assume that MOP results from asymmetric jaw oscillation cycles. Kinematic data were obtained using a Carsten 2D Electromagnetic Articulograph (AG-200 EMA). Jaw oscillations were investigated for duration, velocity, amplitude of the opening vs closing phases, and then statistically analyzed according to different syllable structures. The results on French and Vietnamese bring out differences between the opening and closing phases, showing clear duration and velocity asymmetries. These results are, however, not consistent with previous works on American-English which can provide a phonetic explanation for MOP, based on the properties of the jaw cycles.

1. Introduction

Syllables with an asymmetric structure, i.e. syllables with different onset and coda structures, are frequently found in the world's languages. For example, CV syllables (empty coda), CCVC (complex onset and simple coda), and VC (empty onset) have an asymmetric structure. We call this asymmetric syllable structure 'phonological asymmetry' henceforth. The present study attempts to provide an explanation for this cross-language tendency. Providing the link between phonetic and phonological asymmetries in the structural organization of syllables should allow an explanation of the Maximum Onset Principle (MOP; Kahn 1976; Selkirk 1982; Clements & Keyser 1983; Goldsmith 1990; Blevins 1995). This universal principle of syllabification supports the view that consonants tend to be syllable-initial rather than syllable-final; consonants tend to be attracted in the prevocalic position (thus acting as onsets) rather than in the post-

vocalic position (as codas). MOP is supported by a relatively large number of empirical observations and is widely used in linguistics, psycholinguistics, or in natural language processing. From a theoretical point of view, two frameworks can provide a phonetic base of MOP: the Frame/Content Theory (MacNeilage 1998, 2008) and Articulatory Phonology (Browman & Goldstein 1988, 1990, 1995, 2000).

According to the first theoretical framework, the universal organization of speech is supposed to be a result of jaw oscillation. Indeed, the Frame/Content Theory suggests that a close link exists between the universal CV structure and the cycle of jaw movement: a consonant is produced when the jaw is in a high position whereas a vowel is made when the jaw is in a low position. The cycle of jaw movements thus provides the “Frame” while the segments (consonants and vowels) are said to determine the “Content”. The Frame/Content model also predicts that CV sound combinations favored in natural languages are similar to children’s productions at the stage of babbling, irrespective of the language(s) spoken in their linguistic environment (Davis & MacNeilage 1995, MacNeilage & Davis 2001). These favored combinations called Pure Frames are produced with a single cycle of jaw oscillation without front-back movement of the tongue during consonant-vowel productions. As a result, they are considered simple to produce with a minimum articulatory cost, what should explain their overrepresentation cross-linguistically. However, Rousset (2004), then confirmed by Vallée *et al.* (2009), showed that the proportion of Pure Frames among VC sequences found in the lexicon of 15 languages was significantly higher than the proportion of Pure Frames among CV syllables. This observation suggested a stronger articulatory cohesion between vowel and consonant when the latter is in syllable-final position compared to when it is syllable-initial. The authors thus interpreted such result as evidence in support of an onset vs rime syllable structure, as proposed by those phonologists who argue that syllable structure is rather hierarchical than linear (among others, Klein 1993).

An unsatisfactory feature of the Frame/Content Theory is that it does not provide any clear explanation as to why the preferred structure is consonant-vowel rather than vowel-consonant. Yet the VC structure matches as perfectly as CV with a single jaw cycle, but it represents less than 5% of the syllables found in the world’s languages (Rousset 2004). Another weak point of the Frame/Content Theory which should be noted, is that the CVC structure is the second highest syllable type in the languages of the world, and the highest in some languages (for example, languages of Southeast Asia like

Nyakhur, Thai, Vietnamese, Wa, etc). This syllable type can still contain Pure Frames in it (in onset-nucleus or in nucleus-coda). However, Vallée *et al.* (2009) have found that Pure Frame CVC syllables are underrepresented in the world's languages, since consonant place repetition within a syllable (e.g. /pap/ and /tat/ syllables compared to /pat/or /tap/) are largely avoided.

According to a different perspective, Articulatory Phonology suggests that jaw has no direct role in syllable production, and is considered as merely supporting the actions of others articulators, namely, the lower lip and the tongue. The universal predominance of the CV structure is explained by natural features of consonants, which are “in-phase coupling” with the following vowel. Articulatory gestures for a C and a subsequent V are closely coordinated and compatible with each other without negative mutual interference, because they are realized thanks to independent articulatory mechanisms that allow, at the same time, a synchronized trigger of gestures for C and V. By analyzing gestural overlap, substantially estimated in the time domain by measures of the C-Center Effect (Browman & Goldstein 1988), Articulatory Phonology states that CV combination is “the most stable coupling mode”, and the different intra-syllable positions which are allowed for C (onset or coda) correspond to different configuration modes of gestures (Browman & Goldstein 1995; Byrd 1995). Within a syllable, onset and coda consonants show differences in the temporal organization of articulatory gestures with the vowel. Measurements of the time interval between the consonantal target (or the mean of the consonantal targets in the case of clusters; C-Center) and the vowel target indicate a much tighter and global coordination with the vowel gesture for prevocalic consonants. For syllable-final consonants, the gestural coordination is observed between the first postvocalic-consonant and the vowel. This suggests a global syllable gesture for onsets and a more limited gestural coordination for codas. In contrast to the Frame/Content Theory, the syllable corresponds in Articulatory Phonology to “a gestural constellation”, which is the product of several individual actions corresponding to coarticulated segments. The favored CV combinations are those for which gestures are most naturally in phase and therefore stable (Goldstein *et al.* 2006, Whalen *et al.* 2011) while the reverse VC structure seems to require a more local, less cohesive and less adaptable motor coordination relative to the vocalic gesture. Moreover, in coda position it is expected that the production of the consonant, when anticipated, masks the acoustic realization of the vowel. Consequently, a VC sequence shows much less stability in the

coordination of gestures and shows no C-Center effect. This fact may also explain the vulnerability of consonants in syllable-final position, which are more subject to lenition than consonants in syllable onset. One question that still remains to be investigated is whether the higher proportion of Pure Frames within VC syllables compared to CV syllables can also be explained by the gestural characteristics of the nucleus-coda coordination.

A possible explanation of this apparent nucleus-coda coarticulation can be deduced from the results obtained by Kelso et al. (1985), Gracco (1994), Redford (1999) and Redford & van Donkelaar (2008). Gracco (1994) investigated the opening and closing phases of the jaw cycles in /sV₁CV₂/ sequences, with C=/p b m/, V₁=/i æ/ and V₂=/æ/. In Redford's study, four syllable types (CV/ʃ/, /ʃVC/, /s/CV/ ʃ/, /ʃVC/s/ where V=/i a u/ and C=/p t k/) were chosen to account for differences between the two phases of the jaw cycles in closed syllables (CVC, CCVC and CVCC). From continuous repetitions of five syllable types (CV, VC, CVC, CCV and VCC), Redford & van Donkelaar (2008) attempted to demonstrate that the open-close jaw cycles are consistent with the syllable structure. They were not able, however, to clarify in what extent the jaw cycles influenced the phonetic realization of speech segments and, conversely, in what extent speech segments constrained jaw movements. Kelso and colleagues' study was designed to examine temporal organization of speech by removing as much as possible the individual characteristics of speech segments and preserving the prosodic pattern. The spatiotemporal characteristics of CV forms were analyzed from reiterant speech in which normal syllables of two sentences were replaced by /ba/ or /ma/. They found differences between the opening and closing phases of the jaw cycles at the level of duration, velocity and displacement, and linked them to the dynamic properties of the cycles.

The consistency of results across these four studies with different design, may lead to a phonetic explanation for MOP. By seeking to understand the spatiotemporal organization of speech or seeking to evaluate the impact of biomechanical properties of the jaw on the articulation of segments, these studies revealed an asymmetry between the two phases of the jaw cycle: (i) the raising phase (closing) involves a greater peak of velocity than the lowering phase (opening); (ii) the opening phase has a greater duration than the closing phase; (iii) the lowering phase is articulated with a greater displacement (distance between minimum and maximum opening positions) compared to the raising phase; (iv) the degree of jaw stiffness (estimated by calculating the slope of the correlation between displacement

and velocity) is found to be smaller for the opening phase than for the closing phase. In Gracco's study, the opening and closing phases were found to covary suggesting that oral openings and oral closings are not independent events, as for instance, the covariation found between the jaw opening velocity (for the vowel) and the subsequent closing velocity (for the consonant). Furthermore, Gracco suggested that oral closing depends on oral opening, by the fact that the faster the opening the faster is the closing. A possible explanation given by Gracco is that the oral opening, associated with the vocalic sound, is generally slower because it is involved in shaping resonance and providing the correct VOT, while the closing must be faster to integrate the laryngeal and supralaryngeal combined adjustments.

Following the findings of these previous studies, we assume that if these asymmetries in the properties of the jaw cycle were confirmed in a large amount of languages, then they could be said to account for several tendencies found in the world's languages (see, among others, Rousset 2004): (i) CV rather than VC as the universal canonical syllable; (ii) strong preference for syllable-initial clusters over syllable-final clusters; (iii) complex consonants (e.g. prenasalized, labialized, clicks, affricates, doubly-articulated stops, palatalization) uncommon in coda position; (iv) larger proportion of Pure Frames in nucleus-coda sequences (rime) than in onset-nucleus sequences; (v) cross-linguistically, reduction in the size of phoneme inventories and predominance of consonant lenitions in syllable-final position (constraints on features).

The purpose of the current study is to evaluate whether the jaw cycle plays a role in the understanding of the physical nature of syllables. While previous works (cited above) focused mainly on American English, our ongoing study involves a broader multilingual perspective (French, Polish, Brazilian Portuguese, Ruwund, Tachelhit, and Vietnamese) to find out features of the jaw cycle that could provide a phonetic basis for MOP. The selected languages show various degrees of complexity in their syllable structures and consonant realizations, as well as differences in coda lenition processes and phonological restrictions on codas.

In this paper, we present the first results obtained for an investigation of French and Vietnamese speech. Although only reiterant speech productions of VV sequences are analyzed to date for Vietnamese, the reason why we opted first to compare these two languages needs clarification. Vietnamese is traditionally described as a monosyllabic tone language (Đoàn 1999). Much dissimilarity between French and Vietnamese refer to their lexical and syllabic structures (Tran, 2011). The Vietnamese syllable types are captured by the pat-

tern C1(w)V(C2), the brackets indicating the optional constituents. This pattern implies that the glottal stop which always appears in onset position is phonemic, according to Doãn (1999). Syllabic structure diversity is much more present in French, due to the fact that complex onsets and codas are allowed (C1)(C2)(C3)V(C4)(C5)(C6)(C7) (Rousset, 2004). The most favored syllable types in both languages are CVC and CV. But Vietnamese is predominantly a CVC language while French is mostly a CV language. Like other Southeast-Asian languages, Vietnamese is typologically an isolating CVC language. Therefore, in contrast to French, Vietnamese favors a symmetrical syllable structure and disallows complexity in onset and coda positions.

2. Methodology

2.1. Experimental setup

Tongue, jaw, and lip movements were recorded during a speech production task using the electromagnetic articulography system (EMA) AG200 developed by Carstens®. The acquisition in 2D was conducted at a frequency of 200 Hz, thanks to five coils attached to the articulators (jaw, lower lip, upper lip, tongue tip and tongue body), and two reference coils (attached on the bridge of the nose where the skin movement is minimal and above the upper central incisors) for aligning head movements on the mid-sagittal plane of the subject.

The acoustic signal was recorded via a digital Marantz PMD670 stereo recorder and a directional microphone AKG C1000S, and subsequently digitized at 44.1 kHz.

2.2. Corpus and participants

The preliminary results presented in this paper come from five repetitions of three pseudowords pronounced at a normal speech rate by French speakers:

- /aiaiaiaiai/ (5 jaw cycles)
- /tataitat/ (3 jaw cycles)
- /sasalsas/ (3 jaw cycles)

Regarding Vietnamese, the first results presented here are based on five repetitions of the /aiaiaiaiai/ sequence pronounced by Vietnamese speakers.

These data were extracted from a larger corpus containing other pseudowords and other structures: CV.LVC, CLV.CVLC, CLV.CVCL, CLV.CLVC; C=/p b t d s z/, L=/l/, V=/i a/; the point represents a syllable boundary, and a list of sentences. For Vietnamese, the pseudow-

ords were monosyllabic sequences C_1V and C_1VC_2 , with $C_1=/b\ t\ d\ s\ z/$, $C_2=/p\ t/$ (only $/p/$ and $/t/$ are used as coda consonants in Vietnamese), $V=/i\ a/$ and modal tone B1-D1.

Pseudowords were used in this study because they provide the possibility to avoid as much as possible the influence of both lexical knowledge (frequency or familiarity) and semantic processing. Pseudowords were also used to control phoneme distribution, and smooth any contextual effects in the production of sound sequences. We varied the vowel to explore the effect of the degree of opening on jaw cycle. The consonants are the same that those used in the previous studies cited above. We analyzed firstly alveolar obstruents $/t\ s/$ because they share the highest jaw position, and show the greatest vertical displacements of the jaw when coarticulated with low vowels.

Five female native-speakers of French (CC, EB, LV, ML, SB) aged from 20 to 43 and five female native-speakers of Vietnamese aged from 28 to 35 and coming from Northern Vietnam (HOA, LINH, THANH, HOAI, HUONG) participated in the experiment. The pseudowords were presented one by one on a normal screen placed in front of the speakers. Five repetitions of each trisyllabic pseudoword plus the $/aiaiaiaiai/$ sequence, as well as five repetitions of the sentences, were presented in a random order and in two separated blocks.

Instructions given to the participants before the recording sessions began were to produce each item at a normal speaking rate and to pronounce a clear $/i/$ in the sequence $/aiaiaiaiai/$. This was verified before the recording sessions by asking them to pronounce the $/aiai-aiaiai/$ sequence once and then they were trained with 10 pseudowords (2 for each structure; with $C=/f\ m\ k/$, $L=/R/$, $V=/a\ i/$), presented in a random order. Participants were shown the pseudowords on a computer screen one at a time using a powerpoint presentation with approximately 500ms intervals between slides (presentation was run manually by the experimenter).

The first results presented below are based on 50 CV.CVL. CVC tokens (2 pseudowords * 5 repetitions * 5 speakers) and 50 minus 2 $/aiaiaiaiai/$ tokens (5 repetitions * 5 speakers * 2 languages). Two repetitions of $/aiaiaiaiai/$ by two Vietnamese speakers were disregarded because pronounced $[ajajajaja(j)]$ with a palatal glide instead of the high vowel expected.

2.3. Measures

Using semi-automatic segmentation with EasyAlign and Praat, the pseudowords were extracted and labeled with Trap v.6, a tool developed in GIPSA-lab by C. Savariaux in the Matlab® environment.

The following parameters were measured:

duration of the opening phase (corresponding to jaw lowering) and of the closing phase (corresponding to jaw raising), measured between successive maximum and minimum of the jaw trajectory (vertical component); these values were obtained by detecting the zero-crossing points on the velocity curve;

peak velocity and average velocity of each phase;

amplitude of each phase, which corresponds to the vertical displacement of the jaw, estimated between successive maxima of jaw opening and jaw closing.

2.4. Statistical Analysis

Our goal is to study variations in the response variables (duration, velocity and amplitude of each phase of the jaw cycle) as a function of syllable structure (CV, CVL, CVC, and the sequence of 10 vowels in the /aiaiaiaiai/ stimulus) and, for the /aiaiaiaiai/ stimulus only, as a function of language (French, Vietnamese).

Because our set of data contains repeated measurements made on the same speaker, we opted for linear mixed-effects models, and in order to respect the hypothesis according to which the residues follow a normal rule (condition for the application of mixed models), we chose to transform the response variable in its logarithm (ln, natural logarithm), thus:

$$\log Y_{ijkl} = \mu + \alpha_k + \beta_l + (\alpha\beta)_{kl} + \tau_i + \varepsilon_{ijkl}$$

in which:

Y_{ijkl} : is the j^{th} value of subject i taken by the response variable for modality k of factor A (syllable type), and l of factor B (language);

α_k : is the effect of modality k of factor A on the response variable;

β_l : is the effect of modality l of factor B on the response variable;

$(\alpha\beta)_{kl}$: is the effect of the interaction of modality k of factor A, and modality l of factor B, on the response variable;

τ_i : is the random effect of subject i on the response variable. follows a normal distribution in which is the response variable between the subjects.

ε_{ijkl} : represents the error term and follows a normal distribution.

To analyze differences between the two modalities (i.e. opening vs closing phases), the method of Hothorn et al. (2008) was applied in each modality of the syllable type and language, which allowed us to carry out multiple comparisons of means with the mixed model by ensuring that the risk of the first type related to all simultaneous decision making does not exceed the limit fixed in advance at 5% by adjusting the p-values. The method was applied to the data with the function `glht` of the package `multcomp` of the software R as well as the function `lsmeans` of the package `lsmeans`.

3. Results

3.1. Duration

The multiple comparisons of the average durations between the opening phase and the closing phase for each of the three jaw cycles found in the pseudowords CV.CVL.CVC (C={/t/, /s/}, V=/a/) and the five cycles of the sequence /aiaiaiaiai/ showed significant differences between the two phases, whereby closing is longer than opening (see Table 1, and Figures 1 and 2). An exception is found in the realization of the initial syllable /ta/(t), for which the mean value of opening duration is 7ms greater than the value for closing. However, this difference is not significant. For this syllable, different patterns were observed among the five speakers (see Figure 3 for details), which explain the small and insignificant difference between the mean duration of the two phases.

Table 1. Estimations of mean differences (calculated after log-transformation from the linear model) between closing phase duration (C) and opening phase duration (O) for all the jaw cycles of /aiaiaiaiai/, /tataitat/ and /sasalsas/, with the standard deviation of the differences, the value of the statistics, and the p-value. In brackets are the onsets of the subsequent syllables.

	ESTIMATE	STD. ERROR	Z	P
C-O iai	0.21906	0.04275	5.125	< 0.001 ***
C-O sa.(s)	0.22734	0.06759	3.364	0.00537 **
C-O sal	0.47024	0.06759	6.957	< 0.001 ***
C-O sas	0.46185	0.06759	6.833	< 0.001 ***
C-O ta.(t)	-0.06164	0.06610	-0.932	0.95156
C-O tal	0.31084	0.06610	4.702	< 0.001 ***
C-O tat	0.50349	0.06610	7.617	< 0.001 ***

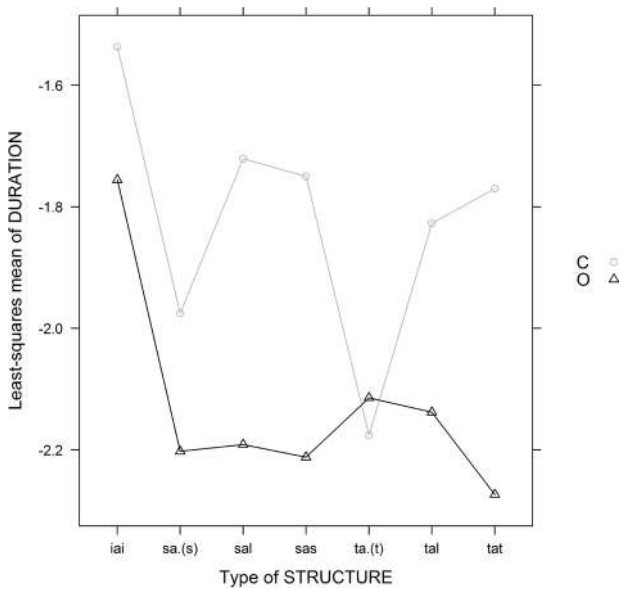


Figure 1. Estimation of average durations of opening and closing phases (log-transformed from the linear model) for each syllable type and /iai/ sequences. Closings are generally longer than openings. Least-Squares Mean is the mean estimated from the linear model.

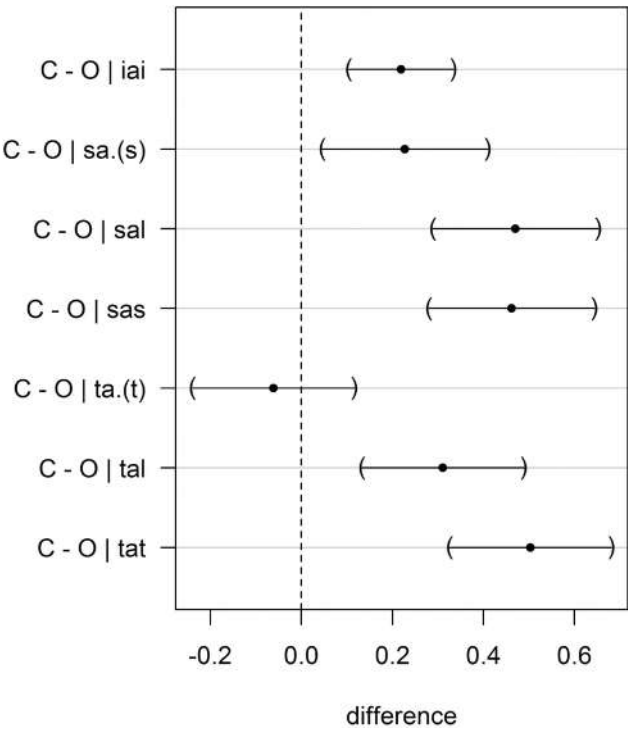


Figure 2. Point estimation of mean difference (from log-transformation) between closing phase (C) and opening phase durations (O) and 95% confidence intervals as a function of sequence type. Note that if the confidence interval includes zero, then the difference is not significant. Intervals are centered on the mean difference values, and their widths depend on variances and sizes. The wider the interval is, the more variability in the sample, the less accurate the estimate of the mean difference.

As can be seen in Figure 1, the raising of the jaw is of longer duration in the case of closed syllables. Against all odds, the longest phase durations (opening as well as closing) are observed in /iai/ sequences, perhaps because only vowels occur in the succession which requires more than one syllable gestures (in the sense of Articulatory Phonology). The comparison between jaw-closing phase durations of /iai/ sequences and that of open CV syllables indicates a shorter closing movement when the following segment is a syllable-initial consonant. The mean durations of jaw-opening phases exhibit less variation than closing phases according to syllable types. A lengthening effect in the pseudoword-final, most probably due to the French iambic accentual pattern, is only observed for the jaw-closing phase towards the alveolar plosive (/tat/).

Figure 2 shows the mean duration differences between the closing and the opening phases. Only for /ta/(t) the difference is not significant. For the other sequences, the closing-to-opening ratio is positive; for the /iai/ sequence, the confidence interval is smaller, indicating less variability and therefore a more stable production compared to the other sequences.

Phase duration differences are in general smaller for the open syllables /ta/(t) and /sa/(s) than for the CVC types (Figure 3). Closing durations are longer for syllables with an offset consonant (on average 170ms vs 126ms). With the exception of /ta/(t), the closing phase is thus consistently longer than the opening phase in the production of French speakers, regardless of syllable structure, type of the consonant, and position in the pseudowords (see Figure 3). Just a slight trend towards the reverse pattern (the opening phase being longer than the closing phase) is observed with ML speaker for /ta/. This unexpected result is not consistent with those of previous studies and especially as in CVC syllables initial C are most often longer than final C (Gracco, 1994; Redford, 1999).

The two phases of the jaw cycle were not related in their timing since no correlations were found between the opening and closing durations.

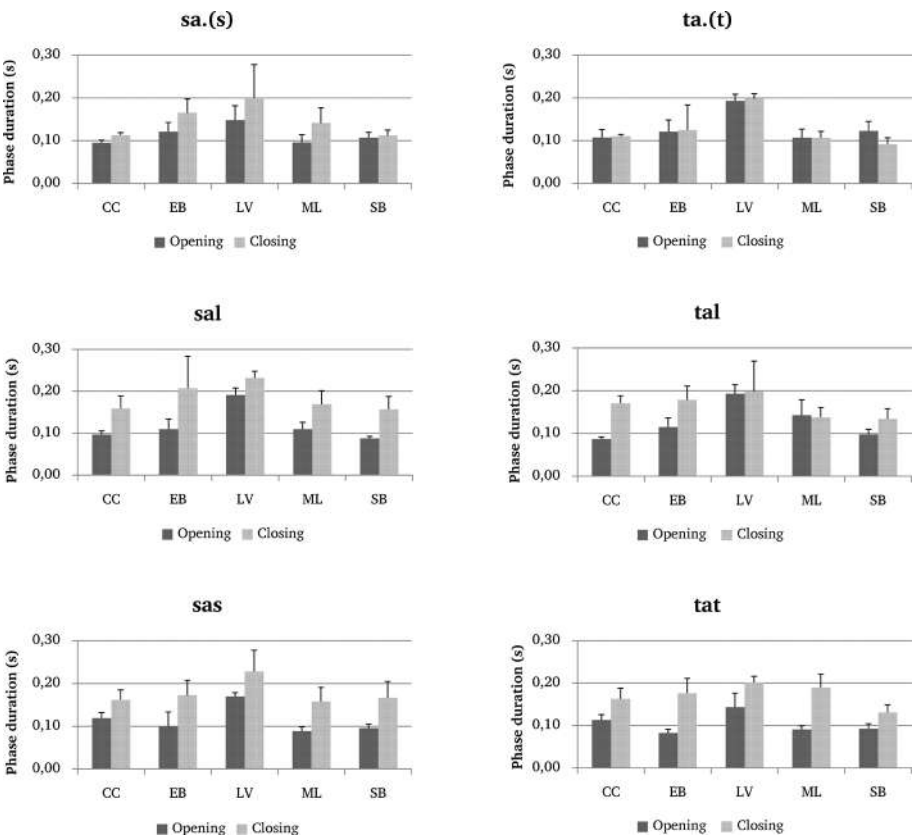


Figure 3. Mean durations of jaw-opening and jaw-closing phases according to syllable types, French female speakers, and both alveolar consonants /t/ and /s/.

3.2. Velocity

As in the case of duration (§3.1), velocity measurements of basic jaw movements reveal an asymmetric pattern. Measurements of the average velocity of opening and closing phases are in most cases significantly different (Table 2, Figure 4), with a faster opening and a slower closing (Figure 5). Only the initial open syllable /ta/(t) does not follow this pattern showing similar average velocity for the two phases; the confidence interval in Figure 5 also indicates a large variability for this type of structure. Again, the smallest confidence interval is observed for /iai/ which is consistent with a more stable production pattern for this type of sequence compared to the others (Kelso

Table 2. Estimations of mean differences (calculated after log-transformation) between closing phase velocity (C) and opening phase velocity (O) for all the jaw cycles measured in /aiaiaiaiai/, /tataitat/ and /sasalsas/, with the standard deviation of the differences, the value of the statistics and the p-value.

	ESTIMATE	STD. ERROR	Z	P
C-O iai	-0.14975	0.05811	-2.577	0.06772
C-O sa.(s)	-0.26517	0.08620	-3.076	0.01458 *
C-O sal	-0.59029	0.08620	-6.848	< 0.001***
C-O sas	-0.44852	0.08620	-5.203	< 0.001***
C-O ta.(t)	0.02152	0.29353	0.073	1.00000
C-O tal	-0.38905	0.08431	-4.615	< 0.001***
C-O tat	-0.33608	0.08431	-3.986	< 0.001***

et al. 1985). Differences between opening and closing are more important for the other syllables, and especially for CVL structures (/tal/, /tat/, /sa/.(s), /sal/ and /sas/).

Looking then at cross-individual variability (Figures 6 and 7), it can be shown that for the /ta/.(t) sequence, two speakers (ML and SB) have an inverse jaw velocity pattern (closing phase faster than opening) compared to the other speakers, while for the /tat/ sequence, only

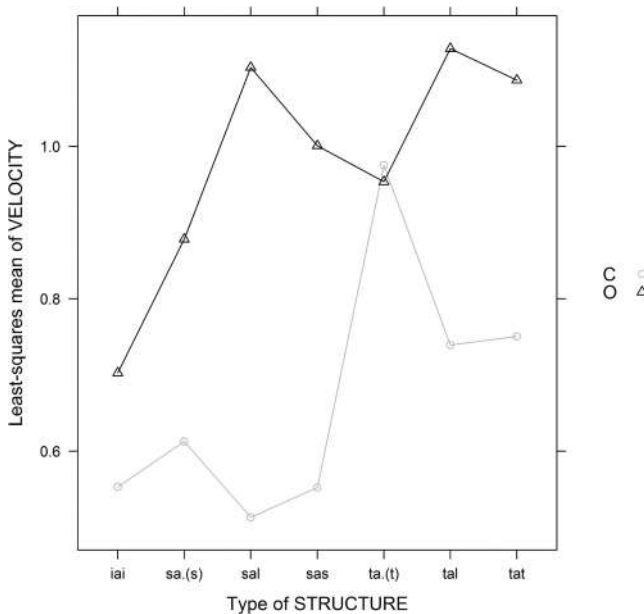


Figure 4. Estimations of the average velocity (log-transformed) of jaw-opening and jaw-closing phases as a function of sequence type. Except for /ta/.(t), the opening phases are faster than the closing phases.

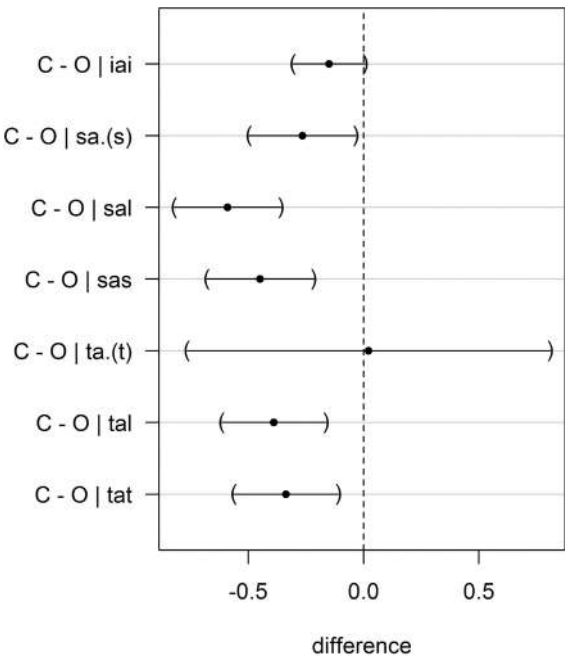


Figure 5. Point estimation of mean average velocity difference (from log-transformation) between jaw-closing (C) and jaw-opening phases (O), and the 95% confidence intervals as a function of sequence type. If the confidence interval include zero, then the difference is not significant.

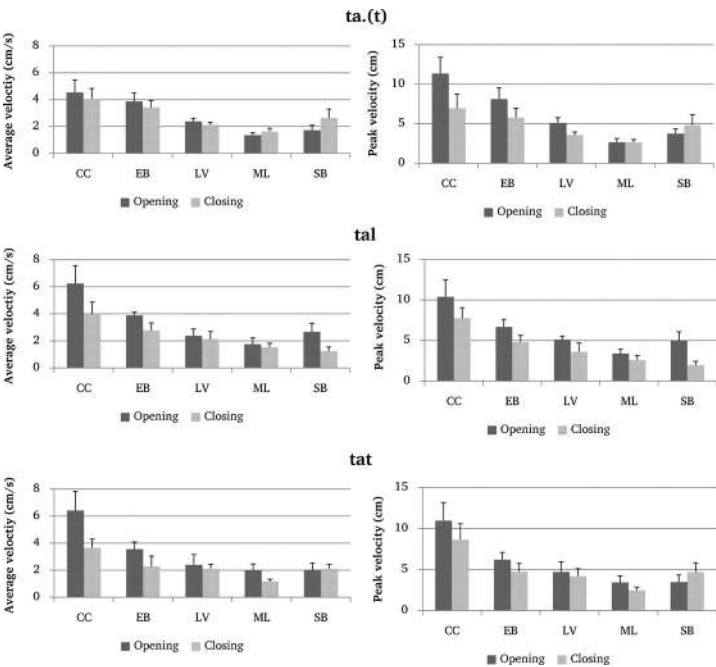


Figure 6. Mean values of average velocity (left) and peak velocity (right) of jaw-opening and jaw-closing phases for the syllables with /t/ across sequence types and speakers.

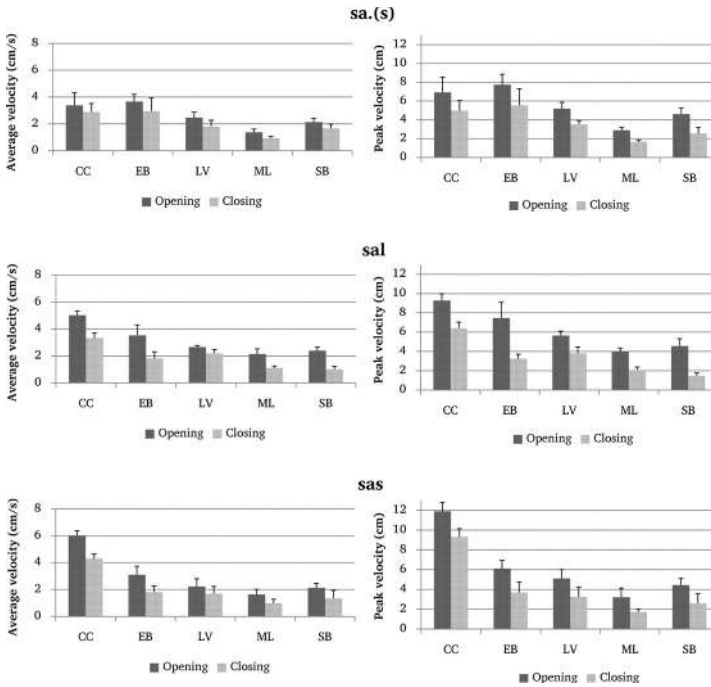


Figure 7. Mean values of average velocity (left) and peak velocity (right) of jaw-opening and jaw-closing phases for the syllables with the fricative consonant /s/ across sequence types and speakers.

one speaker (SB) has a slightly faster closing phase (Figure 6). The results for velocity peaks follow the patterns described for the average velocities. There were however no cross-individual differences for /s/-initial sequences (Figure 7).

For three of the speakers (CC, EB and LV), we found a statistically significant ($p < .05$) positive correlation between opening average velocity and closing average velocity (Figure 8), as well as between opening peak velocity and closing peak velocity for the sequence types (CV, CVL, CVC). This trend was not observed for the two speakers who showed a reversed velocity pattern for the initial syllable /ta/.(t): we found a weak positive correlation for ML and a negative correlation for SB, but statistically insignificant in the two cases ($p > .05$). When /ta/.(t) sequences were removed from the calculation, a statistically significant positive correlation was found between opening average velocity and closing average velocity for ML ($r = .45$, $p < .05$) but not for SB ($r = .27$, $p > .05$); the same observations were made for the velocity peaks.

In the case of the /aiaiaiaiai/ sequences for the five speakers, the r scores were positive, statistically significant and on average higher

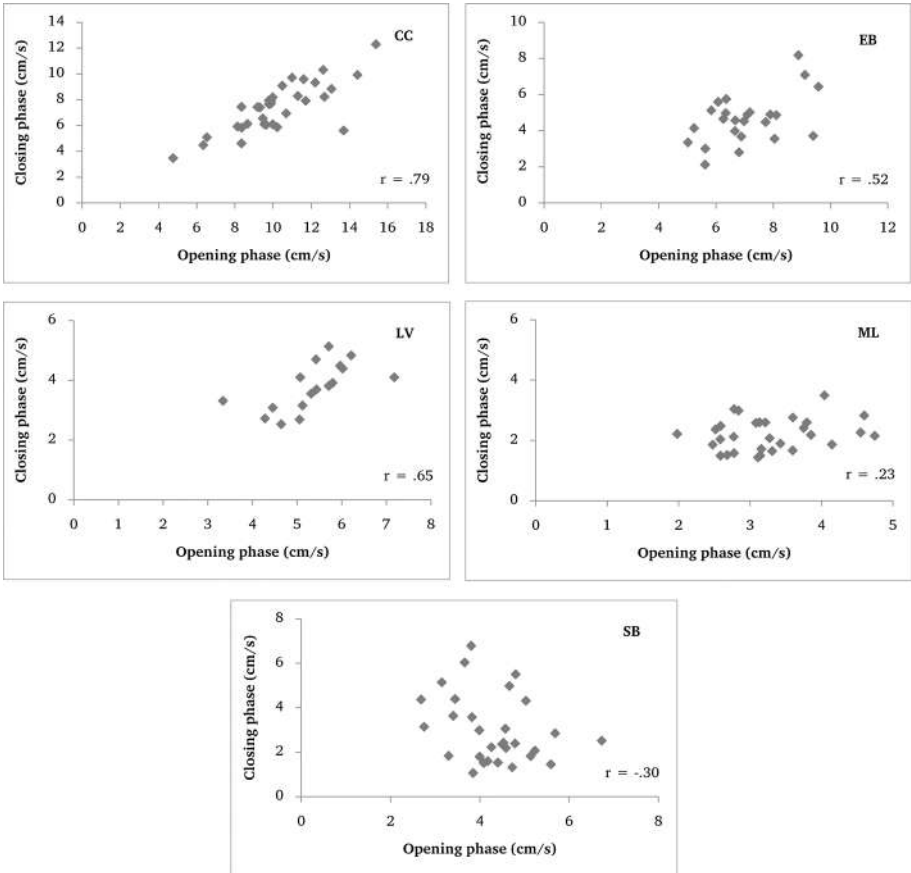


Figure 8. Correlations between jaw-opening and jaw-closing average velocities for each female speaker of French.

than for the other sequences ($r > .73$, $p < .05$). This result therefore suggests a more stable velocity pattern in the production of the vowel sequences.

For all the sequence types, the correlations suggest a relationship between closing phase velocity and opening phase velocity: the faster the opening the faster is the closing.

3.3. Amplitude

The amplitude, estimated as the vertical displacement of the jaw, shows no regular difference pattern between the opening and closing

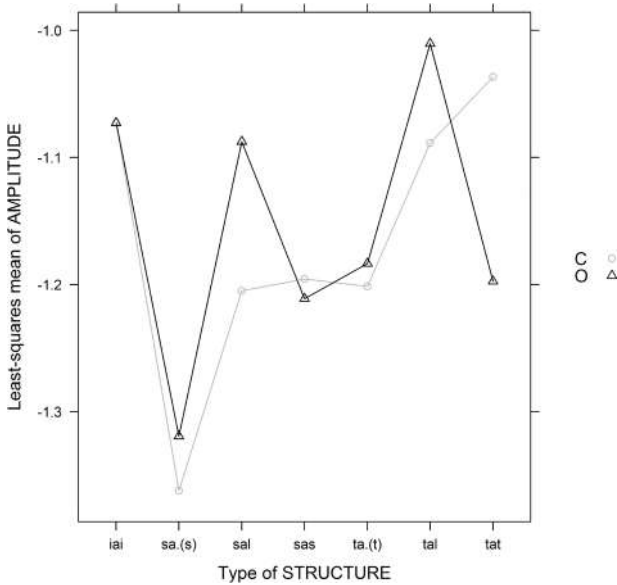


Figure 9. Estimation of average amplitudes of opening and closing phases (log-transformed from the linear model) across sequence types.

phases of the jaw cycles (Figure 9). Figure 10 shows that the differences in the average amplitudes between jaw-opening and jaw-closing phases are not significant for all the sequence types. When the speakers are individually analyzed, no consistency is observed between the differences of phase amplitudes. The same observation was made by Redford (1999) for CVC syllables. This result suggests that the amplitude values depend both on the segments and on the syllable structure types, but more data need to be analyzed to confirm this observation. Across speakers, however, there is a strong positive correlation between jaw-opening and jaw-closing amplitudes in the productions of /tataItat/ and /sasalsas/ sequences (CC: $r = .98$, EB: $r = .96$, LV: $r = .88$, ML: $r = .87$, SB: $r = .65$; for all $p < .05$), as well as in the production of the /aiaiaiaia/ sequences (CC: $r = .97$, LV: $r = .83$, ML: $r = .88$, SB: $r = .73$; for all $p < .05$ and EB: $r = .70$, $p = .05$). The greater the displacement in opening phase the greater is the displacement in closing phase.

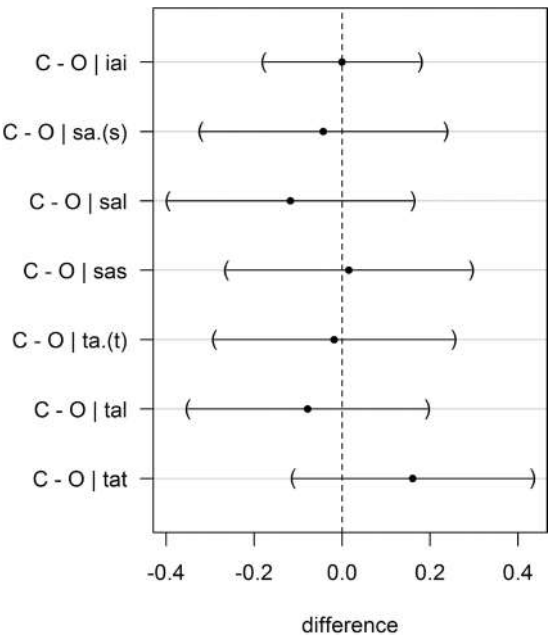


Figure 10. Point estimation of mean amplitude (from log-transformation) between closing phase (C) and opening phase durations (O), and the 95% confidence intervals for each type of jaw cycle. No differences are significant.

3.4. French-Vietnamese comparison

The cross-linguistic comparison only deals with the sequence /aiaiaiaiai/. The overall results of the Vietnamese speakers are similar to those of the French ones. The opening phases are shorter, and with lower amplitude than the closing phases (Figures 11). As for French, differences in durations and average velocities between jaw closing and jaw opening phases are significant (see Table 3 below).

Table 3. Estimations of mean differences of duration, average velocity, and amplitude (calculated after log-transformation) between closing phase (C) and opening phase (O) for all the jaw cycles measured in the sequence /aiaiaiaiai/ produced by the Vietnamese speakers, with the standard deviation of the differences, the value of the statistics and the p-value.

	LANGUAGE	ESTIMATE	STD. ERROR	Z	P
Duration	C-O Fr	0.21906	0.03722	5.886	< 0.001 ***
	C-O Vn	0.16144	0.03067	5.264	< 0.001 ***
Velocity	C-O Fr	-0.14322	0.05726	-2.501	0.0246 *
	C-O Vn	-0.11455	0.04415	-2.594	0.0189 *
Amplitude	C-O Fr	-0.00005	6.297e-02	-0.001	1.000
	C-O Vn	0.0444	5.142e-02	0.864	0.625

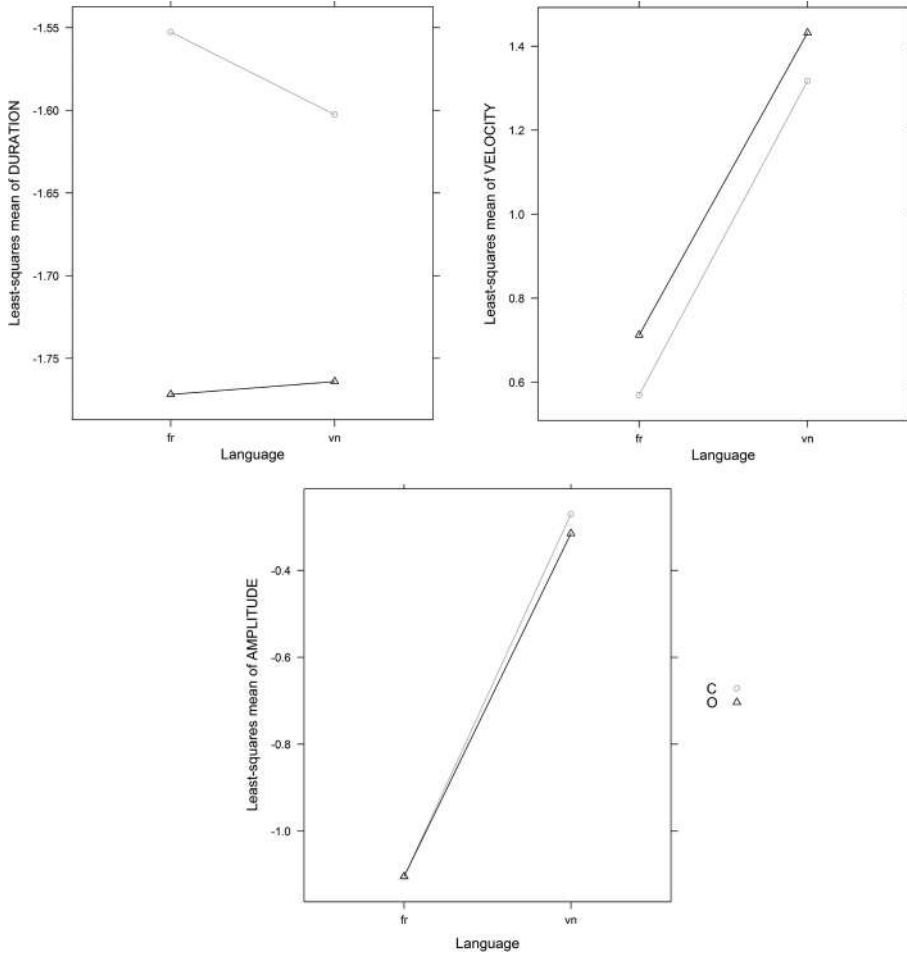


Figure 11. Estimation of (a) mean durations (b) mean velocities (c) mean amplitudes (expressed in log) of jaw-opening and jaw-closing phases for the sequence /aiaiaiaia/ as a function of language.

4. Conclusion

The cross-language preference for syllable-initial over syllable-final consonants in speech, commonly known as the Maximum Onset Principle, is not clearly understood. The present study is part of an ongoing broader multilingual research undertaken to attempt to provide a phonetic basis for MOP. The relationship

between syllable structures and basic open-close jaw movements were investigated within the framework of the Frame/Content Theory and following the findings of several previous studies in which asymmetric patterns in jaw oscillation cycles were observed. We have presented here the preliminary results obtained for French and Vietnamese.

For the two languages, our preliminary results confirm that there are asymmetries in the jaw cycle, with the opening phase (corresponding to the lowering of the jaw) generally shorter, faster, and with smaller amplitude than the closing phase (corresponding to the raising of the jaw). A few changes related to the context (stop vs fricative vs lateral, open vs closed syllables, pseudoword-initial vs final positions) were observed for duration and velocity, but they did not change the pattern of asymmetries. Measurements of the vowel sequences /aiaiaiaiai/ revealed a greater stability of the jaw cycles and stable patterns of asymmetry between the two phases for the duration, velocity, and amplitude parameters, most probably related to the reiteration of successive elements in the sequence (Kelso *et al.* 1985).

Although insignificant statistically, closing phases were found with smaller amplitude for the pseudoword-final syllables, which can be explained by the articulatory release observed in word endings. Several studies have shown that word-initial consonants have a stronger articulation than word-final consonants (for a review see Fougerson 1999; Krakow 1999). The latter are produced with less lingual contact and less linguopalatal pressure (Keating *et al.* 1999, Browman & Goldstein 1995). This word-edge effect in consonant articulation, and more specifically in the peaks of lingual gestures, was much larger for certain consonants and among them /t/, while for others, including /s/, only differences in acoustic duration were observed showing a greater duration for word-initial consonants. From these findings and since jaw supports the action of the tongue, we can expect an effect of word edge on jaw amplitude. However, we need more data to confirm this.

Following Gracco's suggestion that oral opening and oral closing are not independent actions, we calculated the correlations between the jaw opening phase and the subsequent closing phase for the duration, velocity, and amplitude parameters for each speaker. No correlations were found, in our study, between the opening and closing durations of the jaw by the fact that closing durations are different according to the syllable type and are linked to the presence of an offset consonant. Phase duration differences between

opening and closing phases are in general smaller for open syllables. This result was expected since in CV types a single consonant (corresponding to the onset of the subsequent syllable) was articulated during the oral closing-opening movements while in CVC types two heterosyllabic consonants or a single pseudoword-final consonant (with a final lengthening) were articulated. Therefore, from our data, durations of the jaw closing phases seem linked to the syllable types (open vs closed). However, Redford & van Donkelaar (2008) found a lengthening of closing phases when a subsequent CC consonant sequence occurred as an onset (as in V.CCV) rather than as heterosyllabic (as in VC.CV), but this result was not easy to understand for the authors. The lack of correlation for durations of the two phases, observed in our data, is consistent with Gracco's observation that there were no influences of the opening durations on the closing durations of the jaw cycles. In Redford (1999), different patterns of duration in CVC syllable types were observed due to individual differences interpreted at a suprasegmental level (i.e. phrase final lengthening).

For all sequence types, positive correlations were found across speakers between jaw-opening and jaw closing velocities (averages or peaks) and between jaw-opening and jaw-closing amplitudes (displacements). The correlations, with coefficients ranging from $r = 0.73$ to $r = 0.98$, suggest that the two phases of the jaw cycles are strongly related following two tendencies: (1) the faster the opening the faster is the closing; (2) the greater the displacement in the opening phase the greater is the displacement in the closing phase. The fact that these tendencies were consistent across syllable types and across consonant types suggests that they are not necessarily related to the segmental phonological content. These results are consistent with those of Gracco (1994) who concluded that "[...] oral opening and oral closing are not independent events. Rather, oral closing actions are dependent on characteristics of the preceding oral opening action" (p. 16). He suggested that these findings came from biomechanical influences (i.e. muscle activity, tissue elasticity) on oral opening and oral closing movements of speech motor actions.

However, the asymmetrical patterns in the jaw cycle observed in our data with an opening phase generally shorter, faster, and with smaller amplitude than the closing phase contradict the previous works on American English conducted by Kelso *et al.* (1985), Gracco (1994), Redford (1999), Redford & van Donkelaar (2008). As mentioned in the introduction, these authors pointed out a reverse pattern, namely a longer, slower, and wider opening phase that was

assumed to be the result of basic biomechanical properties of the jaw (lowering and raising are not subject to the same constraints). In our study, only the pseudoword-initial sequence /ta/.(t) shows an inverse pattern for both duration and velocity factors in the speech of the two French speakers ML and SB, but statistically insignificant.

This surprising result is not easy to understand. In the introduction, we suggested that consistent patterns of the jaw cycles showing longer and slower opening phases with greater displacements than the closing phases can support the hypothesis of a phonetic basis for MOP by the fact that longer opening and shorter closing durations can better accommodate the articulation of onsets (Gracco 1994). Such asymmetrical open-close movement of the jaw can account for the cross language preferences in syllable types for initial consonants or clusters over final consonants or clusters (also suggested by Redford 1999), as well as the preference for initial complex consonants over final complex consonants, and the larger proportion of Pure Frames in VC (and in vowel-coda of CVC) compared to CV structure. The reverse asymmetric pattern observed in our data is inconsistent with this reasoning. Thus a question to be solved in the future is why are jaw cycle patterns for French and Vietnamese speech exactly opposite to those observed for American English? The consistent pattern across French and Vietnamese excludes for the moment any explanations related to the basic syllable structure of the two languages: French is mostly a CV language while Vietnamese is mostly a CVC. Moreover and in contrast to French, Vietnamese disallows complexity in onset as in coda positions. However, it shows the asymmetrical pattern. Consequently and unlike our first hypothesis, the influence of the characteristics of the jaw cycle on syllable structure (as highlighted by the Frame/Content Theory) is not clarified here and requires further investigation.

However, in our study, the observed asymmetries in duration, velocity and, to a lesser extent (since no significant differences were found), amplitude of jaw cycles are in line with the assumption of Articulatory Phonology that vowels show different phasing patterns with onset and coda consonants. Indeed, a shorter period of time between the point of maximum jaw closing (consonant target) and the point of maximum jaw opening (vowel target) is consistent with a closer phasing of onset-consonant and vowel (in-phase relationship), and also with a high degree of coarticulation, or a synchronous temporal overlap, between C and V gestures. On the other side, a longer and slower jaw closing movement is compatible with a

loose coupling between the nucleus and the following consonant(s). According to Articulatory Phonology, vowels and offset consonants are more time-independent (even in an anti-phase relation), and only the left-edge of a coda is phased with the vowel nucleus, making a coda more variable than an onset (Browman & Goldstein 1988, 1995; Byrd 1995). Within a short time interval it is obvious that intergestural phasing relations are stronger than within a longer time interval. In line with this reasoning, it is not clear how the reverse asymmetric pattern of jaw cycles observed in previous works on American English can be consistent with the basic AP predictions. Gracco (1994) suggested that onset consonants required oral opening and oral closing actions: the lengthening of the opening was helpful for VOT while the shortening of the closing phase supported a tight coordination between laryngeal and supralaryngeal movements. However, this explanation does not appear to account for our findings on French and Vietnamese. It is also inconsistent with Redford & van Donkelaar findings about V.CCV and VC.CV sequences: an unexpected lengthening of the jaw closing phase was observed in the V.CCV sequences. Further works will be required to clearly address this issue.

Results from the present study on jaw-opening and jaw-closing movements in French and Vietnamese speech which show opposite patterns to those observed in previous studies on American English encourage us to continue our multi-speaker and multi-language investigations, and to compare the current results with new data from native speakers of American English. The next step will be to investigate timing relations among lip, tongue and jaw movements, in relation with syllable boundary.

Addresses of the Authors

Nathalie Vallée, GIPSA-lab DPC, UMR 5216, CNRS & Université Grenoble Alpes, BP25 - 38040 Grenoble Cedex 9, France <nathalie.vallee@gipsa-lab.grenoble-inp.fr>

Thi Thuy Hien Tran, GIPSA-lab DPC, UMR 5216, CNRS & Université Grenoble Alpes, BP25 - 38040 Grenoble cedex 9, France <thi-thuy-hien.tran@gipsa-lab.grenoble-inp.fr>

Paolo Mairano, LFSAG, Università di Torino. Via Sant'Ottavio 20, 10124, Turin, Italy <paolomairano@gmail.com>

Silvain Gerber, GIPSA-lab, UMR 5216, CNRS & Université Grenoble Alpes, 11 rue des Mathématiques, BP46 38402 Saint Martin-D'Hères Cedex, France <silvain-gerber@gipsa-lab.grenoble-inp.fr>

Solange Rossato, LIG GETALP, UMR 5217, CNRS & Université Grenoble Alpes, Bât IMAG B, BP53 - 38041 Grenoble Cedex 9, France <solange.rossato@imag.fr>

Acknowledgements

This study is part of the project ANR-10-BLAN-1916 APPSy: Phonetic and Phonological Asymmetries of the Syllable. We would like to thank Christophe Savariaux and Quentin Tura for their valuable assistance in this project. Many thanks also to Myrna Laksman and Steve Huntley, and to the anonymous reviewers for their careful rereading and helpful comments on earlier versions of the paper.

Bibliographical References

- Blevins Juliette 1995. The syllable in Phonological Theory. In Goldsmith John A. (ed.), *Handbook of Phonological Theory*. Oxford: Blackwell Publishers. 206-235.
- Browman Catherine P. & Louis M. Goldstein 1988. Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica* 45. 140-155.
- Browman Catherine P. & Louis M. Goldstein 1990. Gestural Specification Using Dynamically-defined Articulatory Structures. *Journal of Phonetics* 18. 299-320.
- Browman Catherine P. & Louis M. Goldstein 1995. Gestural Syllable Position Effects in American English. In Bell-Berti Fredericka & Lawrence J. Raphael (eds.), *Producing Speech: Contemporary Issues. For Katherine Safford Harris*. New York: AIP Press: 19-33.
- Browman Catherine P. & Louis M. Goldstein 2000. Competing Constraints on Intersegmental Coordination and Self-organization of Phonological Structures. *Bulletin de la Communication Parlée* 5. 25-34.
- Byrd Dani 1995. C-centers Revisited. *Phonetica* 52. 285-306.
- Clements Nick & Samuel J. Keyser 1983. *CV Phonology: A Generative Theory of the Syllable*. Boston: MIT Press.
- Davis Barbara L. & Peter F. MacNeilage 1995. The Articulatory Bases of Babbling. *Journal of Speech and Hearing Research* 38. 1199-1211.
- Đoàn Thín Thút 1999. *Ngữ âm tiếng Việt* (Tr. Vietnamese Phonetics). Hanoi: Nhà xuất bản Đại học Quốc gia Hà Nội (Hanoi National University Publishing).
- Fougeron Cécile 1999. Articulatory properties of initial segments in several prosodic constituents in French. *UCLA Working Papers in Phonetics* 97. 74-99.
- Goldsmith John A. 1990. *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell LTD.
- Goldstein Louis M., Byrd Dani & Elliot L. Saltzman 2006. The Role of

- Vocal Tract Gestural Action Units in Understanding the Evolution of Phonology. In Arbib Michael (ed.), *From action to language: The mirror neuron system*. Cambridge: Cambridge University Press. 215–249
- Gracco Vincent L. 1994. Some Organizational Characteristics of Speech Movement Control. *Journal of Speech and Hearing Research* 37. 4-27.
- Hothorn Torsten, Bertz Frank & Peter Westfall 2008. Simultaneous Inference in General Parametric Models. *Biometrical Journal* 503. 346-363.
- Kahn Daniel 1976. *Syllable-based Generalizations in English Phonology*. Bloomington: Indiana University Linguistics Club.
- Keating Patricia, Wright Richard & Jie Zhang 1999. Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics* 97. 157-173.
- Kelso J.A. Scott, Vatiokiotis-Bateson Eric, Saltzman Elliot L. & Bruce A. Kay 1985. A Qualitative Dynamic Analysis of Reiterant Speech Production: Phase Portraits, Kinematics and Dynamic Modeling. *Journal of the Acoustical Society of America* 77. 266-280.
- Klein Marc 1993. La syllabe comme interface de la production et de la perception phoniques. In Laks Bernard & Marc Plénat (eds.), *De Natura Sonorum: essais de phonologie*. Saint-Denis: Presses Universitaires de Vincennes. 99-141.
- Krakow Rena A. 1999. Physiological Organization of Syllables: A Review. *Journal of Phonetics* 27. 23-54.
- MacNeilage Peter F. 1998. The Frame/Content Theory of Evolution of Speech Production. *Behavioral and Brain Sciences* 21. 499-511.
- MacNeilage Peter F. 2008. *The Origin of Speech*. Oxford: Oxford University Press.
- MacNeilage Peter F. & Barbara L. Davis 2001. Motor Mechanisms in Speech Ontogeny: Phylogenetic, Neurobiological and Linguistic Implications. *Current Opinion in Neurobiology* 11. 696-700.
- Redford Melissa A. 1999. *An Articulatory Basis for the Syllable*. Ph.D. dissertation, The University of Texas, Austin.
- Redford Melissa A. & Paul van Donkelaar 2008. Jaw Cycles and Linguistic Syllables in Adult English. In Davis Barbara L. & Krisztina Zajdo (eds.), *The Syllable in Speech Production: Perspectives on the Frame/Content Theory*. London: Taylor & Francis. 355-76.
- Rousset Isabelle 2004. Structures syllabiques et lexicales des langues du monde: données, typologie, tendances générales et contraintes substantielles. Université Stendhal, Grenoble.
- Selkirk Elisabeth 1982. The Syllable. In van der Hulst Harry & Norval Smith (eds.), *The Structure of Phonological Representations*. Dordrecht: Foris Publications. 337-383.
- Tran Thi Thuy Hien 2011. *Processus d'acquisition des clusters et autres séquences de consonnes en langue seconde: de l'analyse acoustico-perceptive des séquences consonantiques du vietnamien à l'analyse de la perception et production des clusters du français par des apprenants vietnamiens du FLE*. Ph.D. dissertation, Université Stendhal, Grenoble.
- Vallée Nathalie, Rossato Solange & Isabelle Rousset 2009. Favoured Syllabic

- Patterns in the World's Languages and Sensorimotor Constraints. In Pellegrino François, Marsico Egidio, Chitoran Ioana & Christophe Coupé (eds.), *Approaches to Phonological Complexity*. Berlin: Mouton de Gruyter. 111-139.
- Whalen Douglas H., Giulivi Sara, Goldstein Louis M., Nam Hosung & Andrea G. Levitt 2011. Response to MacNeilage and Davis and to Oller. *Language Learning and Development* 73. 243-2.