Representing phonotactics

LabPhon16 Satellite event

June 23, 09:00 - 12:30

Organizers

Ioana Chitoran
Université Paris Diderot & Clillac-ARP

Michela Russo
Université Jean Moulin Lyon 3 & UMR 7023 CNRS Paris 8

Part 1
We will start with individual contributions, which will address the topics of the workshop. The presentations will be grouped together by topic. At the end of each topic there will be a question period.

1. TYPOLOGY AND EVOLUTION **TOTAL TIME: 40'**

9 h – 9 h 10

Shelece Easterday

**Syllable typology and syllable-based typologies:
Findings from the extremes of phonotactic complexity**

Laboratoire Dynamique Du Langage (CNRS & Université de Lyon 2)
shelece.easterday@cnrs.fr

9 h 10 – 9 h 20

Geoffrey Schwartz

Towards a typology of consonant synchronicity

Adam Mickiewicz University in Poznań
geoff@wa.amu.edu.pl

9 h 20 – 9 h 30

Péter Rebrus[1] & Péter Szigetvári[2]

Gradual phonotactics

[1]*Research Institute for Linguistics, Hungarian Academy of Sciences*
[2]*Eötvös Loránd University*
rebrus@nytud.hu ; szigetvari@elte.hu

9 h 30 – 9 h 40: 10 min for questions

2. Perception-production dynamics TOTAL TIME: 65'

9 h 40 – 9 h 50

Matthew Masapollo[1], Jennifer Segawa[1,2], Mona Tong[1], & Frank Guenther[1,3]

Evidence for the consonant cluster as a basic unit of speech motor sequencing

[1]Department of Speech, Language & Hearing Sciences, Boston University
[2] Departments of Neuroscience and Biology, Stonehill College
[3]Department of Biomedical Engineering, Boston University

mmasapol@bu.edu ; jsegawa@stonehill.edu ; monatong44@gmail.com ; guenther@bu.edu

9 h 50 – 10 h

Pierre Hallé

Perceptually repairing illegal clusters: Is there an early faithful representation?

CNRS UMR 7018 Laboratoire de Phonétique et Phonologie, Paris 3, Sorbonne Nouvelle

pierre.halle@univ-paris3.fr

10 h – 10 h 10

Yueh-chin Chang & Feng-fan Hsieh

Variation in responses to conflicting targets in the Mandarin VN rimes

National Tsing Hua University, Taiwan

ycchang@mx.nthu.edu.tw ; ffhsieh@mx.nthu.edu.tw

10 h 10 – 10 h 20

Steven Alcorn and Rajka Smiljanic

Unlearning to perceive and produce epenthetic vowels: the case of L1 Brazilian Portuguese/L2 English sequential bilinguals

The University of Texas at Austin

steven.alcorn@utexas.edu ; rajka@austin.utexas.edu

10 h 20 – 10 h 30

Ela Portnoy & Elinor Payne

The role of phonotactics and lexicality on the perception of intrusive vowels

University of Oxford

elinor.payne@phon.ox.ac.uk

10 h 30 –10 h 45: 15 min for questions

(1h45 into the workshop)

10 h 45 – 11 h 05 **POSTERS and BREAK:** **Total time: 20'**

1. Yoon Mi Oh, Clay Beckner, Jen Hay, Jeanette King

New Zealand Institute of Language, Brain and Behaviour, University of Canterbury
yoonmi.oh@canterbury.ac.nz

Non-Māori speaking New Zealanders show surprisingly sophisticated Māori phonotactic knowledge

2. Harim Kwon[1,2] Ioana Chitoran[2]

[1]*George Mason University*; [2]*Université Paris Diderot*
hkwon20@gmu.edu ; ioana.chitoran@univ-paris-diderot.fr

The adaptation of native clusters with non-native phonetic patterns is task-dependent

3. [1,2]Tom Lentz, [2]Marianne Pouplier, [2]Phil Hoole

[1]*University of Amsterdam;* [2]*IPS Munich, Ludwig Maximilians Universität*

T.O.Lentz@uva.nl; pouplier@phonetik.uni-muenchen.de; hoole@phonetik.uni-muenchen.de

Machine learning shows that clusters representations have a temporal component

(2h05 into the workshop)

3. BEYOND PHONOLOGY TOTAL TIME: 55'

11 h 05 – 11 h 15

Laura Dilley

The role of distal suprasegmental rate and rhythm in phonotactic parsing of speech

Dept. of Communicative Sciences and Disorders, Michigan State University

idilley@msu.edu

11 h 15 – 11 h 25

Boyd Michailovsky

Syllable boundary displacement in Limbu verb-stem alternations

Lacito, CNRS, and LabEx Empirical Foundations of Linguistics, France

boyd.michailovsky@gmail.com

11 h 25 – 11 h 35

Chiara Celata[1], Giulia Bracco[2]

Probabilistic phonotactics in visual word recognition within and across morphological boundaries

[1]*Scuola Normale Superiore di Pisa,* [2]*Università di Salerno*

11 h 35 – 11 h 45

Donald Gong

Grammaticality and lexical statistics in Chinese unnatural phonotactics

University College London (UCL) London, UK

ucjudgo@ucl.ac.uk

11 h 45 – 12 h : 15 min for questions

(3h into the workshop)

Part 2

12 h – 12 h 20 (20 min)
We will split into small working groups, mixing the topics within each group. Each group will propose up to three research questions.

12h 20 – 12h 30 (10 min)
All the participants together will agree on a final list of research questions, considered crucial for new directions of further study.

Syllable typology and syllable-based typologies: findings from the extremes of phonotactic complexity

Shelece Easterday
Laboratoire Dynamique Du Langage (CNRS & Université de Lyon 2)
shelece.easterday@cnrs.fr

Keywords: syllable complexity, phonological typology, holistic typology

Cross-linguistic studies of phonotactics often emphasize implicational generalizations regarding the sequencing of consonants with respect to properties of phonation, place of articulation, manner of articulation, and/or sonority (Greenberg 1965/1978, Morelli 1999, Kreitman 2008, Parker 2012). However, such studies are typically limited to biconsonantal clusters. Thus languages with more complex syllable patterns, as well as those in which consonant sequences do not occur — roughly 14% and 13% of languages, respectively — tend to be underrepresented in typological generalizations about phonotactics. Meanwhile, the property of syllable structure complexity features prominently in many holistic typologies of language. This is the case for typologies centered around phonological holism (cf. Isačenko 1939/1940, Dauer 1983, Auer 1993, Schiering 2007) and those which additionally consider morphosyntactic and semantic properties (cf. Skalička 1979, Fenk-Oczlon & Fenk 2005, 2008). In this talk I present findings which contribute to our understanding of both the typology of complex phonotactics and the interactions between syllable structure complexity and different components of linguistic structure.

In Easterday (2017) I investigated associations between syllable complexity and various linguistic properties in a diversified sample of 100 languages representing four degrees of syllable/phonotactic complexity: Simple, Moderately Complex, and Complex, as defined by Maddieson (2006), and an additional category of Highly Complex. While some associations show a trend across the four categories, the syllable patterns of languages at the extreme ends of the complexity cline, in particular, tend to co-occur with specific sets of phonological and morphosyntactic properties. Languages with Simple phonotactic patterns — that is, canonical syllable structures of (C)V or CV — are characterized by small consonant phoneme inventories, particular kinds of consonant contrasts, low rates of vowel reduction, high rates of consonant allophony, and lower average morpheme/word ratios. Languages with Highly Complex phonotactic patterns — defined as word-marginal sequences of three obstruents or four or more consonants — are characterized by large consonant phoneme inventories, particular kinds of consonant contrasts, high rates of vowel reduction, high rates of morphologically complex clusters, and higher average morpheme/word ratios. The languages in this category also show a great deal of consistency in the reported acoustic properties of their consonant sequences and the distributional properties of consonants within sequences and sequences within syllables. Furthermore, the phonological and morphological properties associated with the category as a whole are more likely to occur in languages in which Highly Complex phonotactic patterns are frequent and relatively unrestricted. These findings suggest that the languages in this category constitute a particularly coherent linguistic type defined by segmental and morphological properties, as well as dynamic processes of sound change, in addition to phonotactic patterns.

Although the data presented here is primarily concerned with the extreme ends of the syllable complexity cline, the results bear relevance for phonotactic typology more generally. In particular, the findings here elaborate upon the properties of high phonotactic complexity and support the idea that syllable patterns can be important defining characteristics of holistic language types.

References

Auer, Peter. 1993. *Is a rhythm-based typology possible? A study on the role of prosody in phonological typology*. Konstanz: Fachgruppe Sprachwissenschaft, Universität Konstanz.

Dauer, Rebecca M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.

Easterday, Shelece. 2017. Highly complex syllable structure: a typological study of its phonological characteristics and diachronic development. (Ph.D. dissertation, University of New Mexico.)

Fenk-Oczlon, Gertraud and August Fenk. 2005. Crosslinguistic correlations between size of syllables, number of cases, and adposition order. In G. Fenk-Oczlon and Ch. Winkler (eds.), *Sprache und Natürlichkeit. Gedenkband für Willi Mayerthaler*, 75-86. Tübingen: Gunther Narr.

Fenk-Oczlon, Gertraud and August Fenk. 2008. Complexity trade-offs between the subsystems of language. In M. Miestamo, K. Sinnemäki, and F. Karlsson (eds.), *Language Complexity: Typology, Contact, Change*, 43-65. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Greenberg, Joseph. 1965/1978. Some generalizations concerning initial and final consonant sequences. *Linguistics* 18: 5-34. (Reprinted in 1978 in Joseph H. Greenberg (ed.), *Universals of Human Language, vol. 2: Phonology*. Stanford, CA: Stanford University Press.)

Isačenko, A. V. 1939/1940. Versuch einer Typologie der slavischen Sprachen [Attempt at a typology of the Slavic languages]. *Linguistica Slovaca* 1: 64-76.

Kreitman, Rina. 2008. The phonetics and phonology of onset clusters: the case of Modern Hebrew. (Ph.D. dissertation, Cornell University).

Maddieson, Ian. 2006. Correlating phonological complexity: data and validation. *Linguistic Typology* 10(1): 106-123.

Morelli, Frida. 1999. The phonotactics and phonology of obstruent clusters of Optimality Theory. (Ph.D. dissertation, University of Maryland).

Parker, Steve. 2012. Sonority distance vs. sonority dispersion–a typological survey. In Parker, Steve (Ed.) *The sonority controversy*. Berlin/Boston: De Gruyter Mouton. 101-165.

Schiering, René. 2007. The phonological basis of linguistic rhythm: cross-linguistic data and diachronic interpretation. *Sprachtypologie und Universalienforschung* 60(4): 337-359.

Skalička, Vladimír. 1979. Typologische Studien. Braunschweig: Vieweg.

Towards a typology of consonant synchronicity

Geoffrey Schwartz (geoff@wa.amu.edu.pl)
Adam Mickiewicz University in Poznań

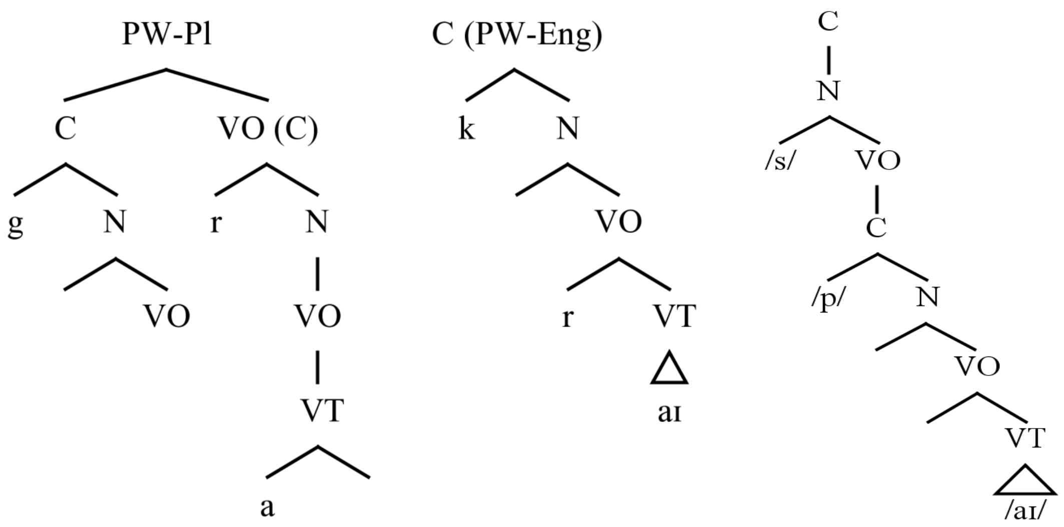
A comparative EMA study of Polish and Tashlhiyt Berber [1] reveals conflicting evidence with regard to the syllabic organization of consonant clusters in Polish. On the one hand, unlike in Tashliyt, right-edge-to-anchor variability was significantly greater than center-to-anchor variability, a finding that suggests 'complex' onset organization in line with established phonetic heuristics [2]. On the other hand, target-to-target lags were notably greater in Polish than in Tashlhiyt. That is, Polish clusters exhibited less synchronicity in cluster timing. If the consonants in the Polish clusters reflected a single 'onset' constituent, as is suggested by the 'complex' onset interpretation, this latter finding is surprising. We should expect greater phonetic cohesion within a single prosodic constituent (complex 'onsets' in Polish) than in a consonant sequence spanning multiple constituents (simplex 'onsets' in Tashlhiyt).

Additional evidence of asynchronous cluster organization in Polish has been observed in the form of numerous intrusive vocoids in onset clusters of various types [3], [4] and asynchronicity in the production of the so-called 'palatalized' stops [d^j] [b^j], which are in fact stop-glide clusters [5]. Beyond this, there is phonological evidence to suggest that 'onset' clusters in Polish bear prosodic weight [6], [7]. CV content words in Polish are prosodically sub-minimal – there are no nouns in the language that have this shape – while CCV words are perfectly well formed. These facts suggest that the first consonant of a cluster in Polish bears a certain degree of prosodic autonomy, resulting in a lack of synchronicity in cluster production, which appears to be an inherent aspect of Polish phonology. The question that remains is how cluster synchronicity may be encoded in phonological representations.

The Onset Prominence (OP) framework [7], [8] offers tools for the representation of three different degrees of cluster synchronicity. Consonant sequences may be absorbed at the same representational level, in which case their articulation is tightly coordinated. This configuration obtains in 'rising sonority' clusters in languages such as English – synchronous articulatory coordination is evident in processes such as approximant devoicing (e.g. *clear*), TR affrication (*try*) and coalescence of /tj/ and /dj/ (*tune*). Clusters may be adjoined at a higher level, in which case they should be asynchronous, and act as if they are separate prosodic units. This is posited for Polish. Alternatively, consonants may be 'submerged', i.e. joined into a single structural constituent, but at different representational levels, yielding an intermediate level of phonetic cohesion. This configuration is posited for non-TR onsets in English, and all clusters in Tashlhiyt. The three configurations are shown in (1). On the left we see an adjoined /gr/ cluster in the Polish word *gra* 'game'. In the center we see an absorbed /kr/ cluster in English *cry*. On the right we see a submerged /sp/ cluster in English *spy*. Crucially, these configurations are the products of independently motivated mechanisms in the OP representational system.

In most EMA research, the organization of 'onsets' is computed with respect to anchors housed later in the syllable. Researchers have varied in their choice of anchor. Some opt to use landmarks associated with a post-vocalic consonant, while others calculate articulatory coordination with respect to a vocalic 'nucleus'. In some cases [2], anchor choice has been found to affect findings with regard to the simplex vs. complex onset hypothesis. The OP approach alleviates this problem by limiting its predictions to target-to-target lag in consonant sequences. Since relative to 'onsets', vowels and post-vocalic consonants play a minimal role in determining phonological constituency in the OP system, they are predicted to be somewhat unreliable as reference points for syllable structure. Rather, cluster synchronicity is a function of timing relations between consonants only.

(1) From left: OP representations for Polish *gra* 'game', English *cry*, and English *spy*



[1] Hermes, A., Mücke, D & Auris, B. 2017. The variability of syllable patterns in Tashlhiyt Berber & Polish. *Journal of Phonetics* 64, 127-144.

[2] Shaw, J. A., Gafos, A., Hoole, P., & Zeroual, C. 2009. Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability in articulation. *Phonology* 26, 187–215.

[3] Cieślak, J. 2015. English Polish differences in the organization of consonant clusters: a study in phonetic code-switching. MA Thesis, Adam Mickiewicz University in Poznań.

[4] Dłuska, M. 1986. Fonetyka polska. Warszawa: Państwowe Wydawnictwo Naukowe.

[5] Swięciński, R. 2012. Acoustic aspects of palatalization in English and Polish. PhD Thesis, UCMS Lublin.

[6] Comrie, B. 2000. Irregular stress in Polish and Macedonian. *International Review of Slavic Linguistics and Poetics* 1, 227-240.

[7] Schwartz, G. 2016. On the evolution of prosodic boundaries – parameter settings for Polish and English. *Lingua* 171, 37-74.

[8] Schwartz, G. 2015. Who needs a nucleus? Tashlhiyt Berber syllabification within the Onset Prominence representational environment. *Poznań Studies in Contemporary Linguistics* 51 (2), 247-290.

Gradual phonotactics

Péter Rebrus <rebrus@nytud.hu>, Research Institute for Linguistics, Hungarian Academy of Sciences
& Péter Szigetvári <szigetvari@elte.hu>, Eötvös Loránd University

Syllable structure is managed not only by sonority, but also by the place of adjacent consonants. It is odd that place sharing is usually, but not always advantageous: while homorganic nasal/liquid+plosive clusters are preferred, heterorganic plosive+plosive clusters may occur without homorganic ones (ie geminates). Besides addressing this asymmetry, we show that phonotactics is more fine grained than what could be captured by a categorical device like syllable structure.

The distribution of consonants is freer before a vowel than before a consonant or word finally. Hooper (1976) and Murray & Vennemann (1983) have examined preferences for CCs occurring where syllables meet. Constraints on preconsonantal Cs have been identified by Itô (1986), Goldsmith (1990), etc. Much work has been done on explaining these preferences in Government Phonology (Kaye & al. 1990, Harris 1990, Charette 1992, Harris 1997, etc), as well as from the viewpoint of perception by, eg, Steriade (1999). (1) shows monomorphemic, intervocalic plosive-final cluster types and their accessibility in languages. CC types can be arranged in an implicational hierarchy, but the accessibility of geminates is independent of the other types (TT = geminate, NT = homorganic nasal + plosive, RT = liquid + plosive, ST = fricative + plosive, PT = heterorganic plosive + plosive, MT = heterorganic nasal + plosive). The generalizations extend to other intervocalic and word-final clusters, (1). There may be great differences between CCs: TT and NT are homorganic, PT and MT are heterorganic, RT and ST may be either. The markedness of CCs depends on homorganicity and the coronals involved (eg, **rt, lt, st** are less marked than **rp, lp, sk**, respectively) and other clusters also often differ in markedness (eg, **pt, mt** are less marked than **tp, np**). Even if a language can access a given cluster type, it may not be able to access all clusters of that type, because of the markedness differences above. In (2), we have counted the ratio of available monomorphemic clusters of all possible clusters in each type in Hungarian (1 = all, 0 = no CCs of the type are well-formed, incomplete types are shaded). The further right a cluster type is in the chart, the smaller the accessibility of the clusters belonging to that type. The greater phonotactic freedom of intervocalic vs word-final clusters, and of nouns vs verbs is also visible.

Our analysis does not presuppose abstract entities (like “syllable” or “coda”) or even skeletal positions, consonantal sequence types are described directly. The data in the charts can be interpreted as a complexity hierarchy of the phonological constructions accessible in the given language. Complexity is measured in the amount of independent information in the coda: eg, none in TT (the two Cs are identical), only nasality in NT. Hence the hierarchy does not faithfully follow the sonority hierarchy. Other CC types contain more and more additional information, further place and manner features. Thus homorganic clusters are always less complex than heteroganic clusters within a type. A CC is well-formed in a language if its complexity is between the minimally and the maximally complex constructions. The CC construction of minimal complexity is TT in languages with geminates, NT in others without. “Lower level” (segmental) constructions not available in the language may impose further constraints: Eastern Ojibwe, for example, lacks liquids, hence the RT construction is unavailable, although ST is available and should imply RT. Inclomplete CC types — which lack some of the potential clusters — exhibit subregularities based on homorganicity and coronality. These subregularities may be expressed by similar complexity hierarchies.

Segmental complexity forms a hierarchy very similar to cluster complexity: the availability of a phonemic glottal stop, **h**, or **ə** (segments of minimal complexity) is independent of the availability of other, more complex segments (cf geminates and other CCs). In fact, in the case of some complex segments there is no difference between the two hierarchies: the prenasalised stop $^{\mathbf{n}}$**d** cannot contrast with the **nd** cluster. Likewise falling diphthongs and glide+C clusters (**a**$^{\mathrm{w}}$**t** vs. **awt**) need not be treated separately.

(1)

	TT	NT	RT	ST	PT	MT	example (intervocalically)
0							Hawaii (Maddieson 2013)
1		↔					Manam (Piggott 1999)
1+	←	→					Japanese (Prince 1984), Pali (Zec 1998)
2		←	→				Diola Fogny (Piggott 1999)
2+	←		→				Sidamo (Gouskova 2004)
3		←		→			Basque (Egurtzegi 2013)
3+	←			→			Italian (Krämer 2009)
4		←			→		Spanish (Hualde 2014)
4+	←				→		Hungarian (Siptár & Törkenczy 2000)
5		←				→	Kashmiri (Wali & Koul 1997)
5+	←					→	Hindi (Kachru 2006)

(2)

	TT 6	**NT** 6	**RT** 12	**ST** 24	**PT** 30	**MT** 15	**types** number of all potential CCs in Hungarian		
V_V	1	1	1	.50	.40	0	nouns	voiceless	well-formedness ratios
V_#	1	1	.92	.21	.13	0			
	.17	.17	.17	.08	0	0	verbs		
V_V	1	1	.75	.29	.07	0	nouns	voiced	
V_#	1	.67	.50	.08	.03	0			
	.50	.33	.17	.04	0	0	verbs		

References

Charette, Monik. 1992. Mongolian and Polish meet government licensing. SOAS WPLP 2: 275–291. ❖ **Egurtzegi,** Ander. 2013. Phonetics and phonology. In Martinez Areta Mikel (ed.) Basque and Proto-Basque: Language-internal and typological approaches to linguistic reconstruction. Frankfurt a. M.: Peter Lang. 119–172. ❖ **Gouskova,** Maria. 2004. Relational hierarchies in Optimality Theory: the case of syllable contact. Phonology 21: 201–250. ❖ **Harris,** John. 1990. Segmental complexity and phonological government. Phonology 7: 255–300. ❖ **Harris,** John. 1997. Licensing Inheritance: an integrated theory of neutralisation. Phonology 14: 315–370. ❖ **Hooper,** Joan Bybee. 1976. An introduction to natural generative phonology. NY: Academic Press. ❖ **Hualde,** José Ignacio. 2014. Los sonidos del español. Cambridge: CUP. ❖ **Goldsmith,** John A. 1990. Autosegmental and metrical phonology. Oxford: Blackwell. ❖ **Itô,** Junko. 1986. Syllable theory in Prosodic Phonology. PhD diss., UMass, Amherst. ❖ **Kachru,** Yamuna. 2006. Hindi. Amsterdam: JBenjamins. ❖ **Kaye,** Jonathan, Jean **Lowenstamm,** and Jean-Roger **Vergnaud.** 1990. Constituent structure and government in phonology. Phonology 7: 193–232. ❖ **Krämer,** Martin. 2009. The phonology of Italian. Oxford: OUP. ❖ **Maddieson, Murray,** Robert W. and Theo **Vennemann.** 1983. Sound change and syllable structure in Germanic Phonology. Lang. 59: 514–528. ❖ **Piggott,** Glyne L. 1999. At the right edge of words. The Ling. Rev. 16: 143–185. ❖ **Prince,** Alan S. 1984. Phonology with tiers. In Mark Aronoff and R. T. Oehrle (eds.) Language sound structure. Cambridge, Mass.: MIT Press. 234–244. ❖ **Siptár,** Péter and Miklós **Törkenczy.** 2000. The phonology of Hungarian. Oxford: OUP. ❖ **Steriade,** Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In Osamu Fujimura, Brian D. Joseph, and Bohumil Palek (eds.) Proceedings of the 1998 Linguistics and Phonetics Conference: Item Order in Language and Speech, Prague: Karolinum Press. 205–245. ❖ **Wali,** Kashi and Omkar N. **Koul.** 1997. Kashmiri: A cognitive-descriptive grammar. London: Routledge. ❖ **Zec,** Draga. 1998. Coda constraints and conditions on moraic projection. WP of the Cornell Phonetics Laboratory 12: 255–303.

Evidence for the consonant cluster as a basic unit of speech motor sequencing

Matthew Masapollo[1], Jennifer Segawa[1,2], Mona Tong[1], & Frank Guenther[1,3]
[1]*Department of Speech, Language & Hearing Sciences, Boston University;*
[2] *Departments of Neuroscience and Biology, Stonehill College;*
[3]*Department of Biomedical Engineering, Boston University*
mmasapol@bu.edu; jsegawa@stonehill.edu; monatong44@gmail.com; guenther@bu.edu

Fluent speech involves rapid sequencing and initiation of motor programs for the phonological units that make up an utterance. The exact nature of these motor sequencing units remains unclear (see Fig. 1). While many researchers have posited basic units that are gestural or phonemic ([1], [2]), others have pointed to larger units, which we will call sub-syllabic constituents (SSCs), that can contain multiple phonemes ([3]), such as syllable onsets, nuclei, and codas. Still others have posited that optimized motor programs exist for entire syllables ([4], [5]). In a prior study from our laboratory [6], English speakers were trained to produce novel phoneme sequences (monosyllabic CCVCC pseudowords) with consonant clusters that were phonotactically illegal in English but legal in other languages. Two days of practice led to measureable performance gains (e.g., reduced utterance durations) for these novel pseudowords, and it was postulated that these gains were primarily due to learning of motor "chunks" for the new consonant clusters, consistent with the view that the units of speech motor sequencing are SSCs. However, the design of that study could not distinguish this possibility from the possibility that the newly learned chunks were syllable-sized.

To address this issue, the current study investigated whether the learning gains found in [6] were specific to trained syllables or whether they generalized to novel syllables containing the newly learned consonant clusters. Generalization to untrained syllables would indicate that the learned motor chunks were smaller than the full syllable. We also compared the learning of novel pseudowords involving phonotactically legal English consonant clusters (e.g., 'flisk') to learning of novel pseudowords with novel (illegal) consonant clusters (e.g., 'gvasf'). If the consonant cluster is the motor chunk learned during training, pseudowords with illegal clusters should show performance gains with practice, whereas pseudowords involving legal clusters should show no gain as these clusters are already well-learned from prior linguistic experience.

We found that the illegal sequences were produced faster and with fewer errors over the two-day period, indicating that speech motor sequence learning occurred. In contrast, we found no significant behavioral gains for the legal sequences. Speakers started out at near-ceiling performance for the production of the legal sequences, presumably because they could produce these sequences by concatenating existing motor programs for native clusters in their production repertoire. Critically, speakers were also faster and more accurate at producing the novel illegal sequences with clusters that occurred in the learned illegal sequences. Moreover, this advantage for producing the previously learned illegal clusters fully generalized to vowel contexts not included in the training stimuli, indicating that learning gains were not specific to entire syllables; instead, once a novel cluster was learned, it could be efficiently produced in new sequences. Collectively, these findings indicate that, at some level, the speech production process entails learning and executing optimized sequences of vocal tract movements that correspond to phonological units smaller than an entire syllable but larger than an individual phoneme.

Keywords: speech production; phonotactics; sub-syllabic constituents; speech sound sequencing; speech motor learning
Theme: Perception-production dynamics

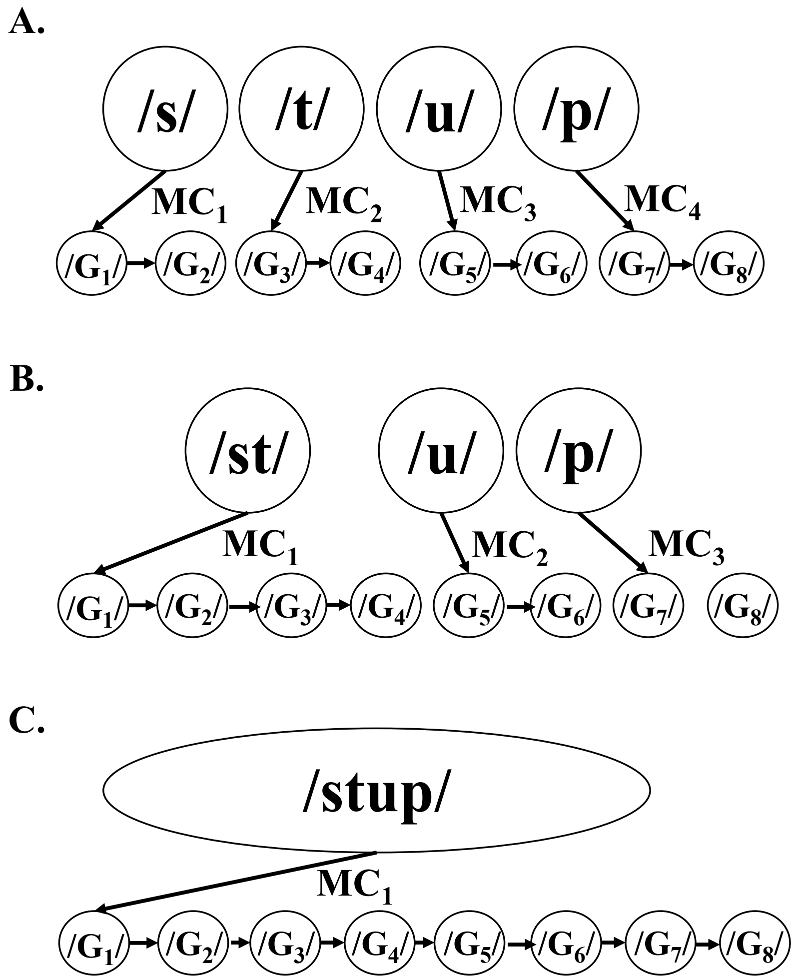


Figure 1. *Three possible accounts of the motor* "*chunks*" *underlying the production of the CCVC sequence* "*stoop.*" *(A.) Four separate motor chunks, one for each individual phoneme. (B.) Three separate chunks, one for each SSC (onset, nucleus, coda). (C.) A single motor chunk for producing the entire syllable. See text for further details. G = gesture, MC = motor chunk.*

[1] Browman, C.P. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica*, 49, 155-180.

[2] Guenther, F.H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, 102 (3), 594 – 621.

[3] Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In: MacNeilage, P. (Ed.), The Production of Speech. Springer-Verlag, New York, pp. 109 – 136

[4] Levelt, W.J., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22 (1), 1 – 38.

[5] Guenther, F.H. (2016). *Neural Control of Speech*. Cambridge, MA: MIT Press.

[6] Segawa, J., Tourville, J.A., Beal, D.S., & Guenther, F.H. (2015). The neural correlates of speech motor sequence learning. *Journal of Cognitive Neuroscience*, 27(4), 819-831.

Perceptually repairing illegal clusters: Is there an early faithful representation?

Pierre Hallé
CNRS UMR 7018 *Laboratoire de Phonétique et Phonologie* LLP/ Paris 3, Sorbonne Nouvelle
pierre.halle@univ-paris3.fr

Many Laboratory Phonology studies, especially those using the OT framework, propose accounts of "misperception" of speech inputs that do not follow the phonological system of the listener's native language, whereby the speech input *in its veridical form* is converted into an acceptable output with respect to the listener's language phonological constraints ([1], [2]). This approach implicitly suggests that listeners' perception unfolds from an initial stage of "faithful" perception in terms of the input language's phonology. Paradis and colleagues (e.g., [3]) have made a similar assumption in their research on loanword adaptations.

The goal of this short study is to review some arguments from the psycholinguistic literature supporting or not the view that the very initial stage of listeners' speech perception is faithful to the phonetic content of the input speech. Sections (a-b) review arguments supporting the veridical perception view, whereas (c-d) review contrary evidence.

(a) Hallé and colleagues ([4]) showed that French listeners most often perceive [tl, dl] as (French) /kl, gl/. This finding has been replicated in [5] using Hebrew /tl, dl/ stimuli: native speakers of French or English, which both ban */tl/ and */dl/ word-initially, perceptually repaired these stimuli as /kl/ and /gl/, whereas native speakers of Hebrew, which allows #/tl, dl/, perceived /tl, dl/ faithfully. [4] included a "phonetic gating" experiment, in which French participants had to transcribe fragments of increasing duration from items such as *tlobda*. The shortest fragments only contained the item-initial stop closure and release burst and the longest ones extended up to the first vowel. Participants' transcriptions for the shortest gates were *faithful* to the speaker's intention in that, for example, they responded "t" vs. "k/c" for the shortest gates of /tlabod/ vs. /klabod/. (They were switching to "kla/cla" transcriptions for the longest gates.) Although the phonetic gating experiment in [4] was intended to control for the speaker's productions at the phonetic level, the authors proposed that the phonetic gating data reflected "snapshots" of listeners' perception as it unfolds over time. However, it can easily be argued that listeners' transcriptions of short fragments might not reflect the time course of perception of the corresponding complete utterances.

(b) Breen et al. [6] conducted an intramodal audio–audio priming experiment with, in particular, /gla/ targets preceded by /gla/, /kla/, or /dla/ primes (identical, different, and critical priming conditions). English listeners had to rate prime–target similarity and their ERPs were collected for the target stimuli. On the behavioral side, they rated /dla/–/gla/ pairs as very similar, contrary to /kla/–/gla/ pairs, thereby confirming the */dl, tl/ to /gl, kl/ phonotactic repair. However, the ERP data rather yielded similar patterns for /dla/–/gla/ and /kla/–/gla/, both differing from the pattern obtained for /gla/–/gla/, suggesting priming (in the form of reduced positivity in the 200-350 ms range) in the latter but not in the former priming conditions. The authors concluded that there should be a perception stage whereby /dla/ is perceived as different from /gla/, that is, presumably, is perceived "veridically."

(c) Dehaene-Lambertz et al. [7] used a straightforward cross-linguistic design, whereby the behavioral and ERP responses to speech contrasts such as /igumo/–/igmo/, legal in French but not in Japanese, were directly tested. French listeners showed a clear MMN response to the /igumo/-/igmo/ contrast. This is the earliest ERP component reflecting presumably non-conscious detection of a phonetic difference. Japanese listeners showed no sign of such an early ERP response, suggesting that phonotactic repair is not preceded by faithful perception.

(d) We finally present some new ERP data on the /tl/-/kl/ contrast, with young infant and adult French listeners. We used a similar habituation-dishabituation paradigm to that in [7], with three habituation precursors followed by a dishabituation target. The target was always /tla/ and the precursors could be /tla/ (no-change), /pla/ (clear-change), or /kla/ (critical-change). The ERP data, obtained in a passive listening condition, show that French 7-month-olds but not adults respond to the /kla/-/tla/ contrast, suggesting that the robust deafness to this contrast is learned. After French listeners have attuned to their native language, there is no sign, in their ERP data, of a response to this contrast (Figure 1).

We further discuss the issue of a faithful perception stage, considering the possibility that this stage has become integrated with phonotactic repair within a language-specific first stage of perception. Indirect evidence may be the increased time-cost found in some studies (e.g., [8]) during the non-conscious processing of phonotactically illegal inputs.

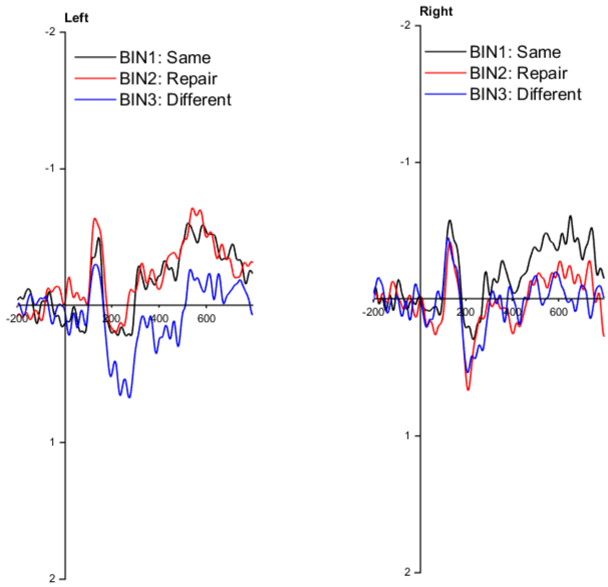


Figure 1. *ERP responses to /tla/–/tla/, /kla/–/tla/, and /pla/–/tla/, pre-central frontal sites, 1st half of experimental session. On the left sites, /tla/–/tla/ and /kla/–/tla/ did not differ, (repair of */tl/ into /kl/), whereas /pla/–/tla/ induced a P300-like response to change.*

[1] Fleischhacker, H. 2001. Cluster-dependent epenthesis asymmetries. *UCLA working papers in linguistics*, 7, 71-116.

[2] Lentz, T., & Kager, R. 2015. Categorical phonotactic knowledge filters second language input, but probabilistic phonotactic knowledge can still be acquired. *Language and Speech*, 58, 387-413.

[3] Paradis, C., & LaCharité, D. 1997. Preservation and minimality in loanword adaptation. *Journal of linguistics*, 33, 379-430.

[4] Hallé, P., Segui, J., Frauenfelder, U., & Meunier, C. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation?. *Journal of experimental psychology: Human perception and performance*, 24, 592-608.

[5] Hallé, P., & Best, C. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/clusters. *The Journal of the Acoustical Society of America*, 121, 2899-2914.

[6] Breen, M., Kingston, J., & Sanders, L. 2013. Perceptual representations of phonotactically illegal syllables. *Attention, Perception, and Psychophysics*, 75, 101-120.

[7] Dehaene-Lambertz, G., Dupoux, E., & Gout, A. 2000. Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience*, 12, 635-647.

[8] Hallé, P., Dominguez, A., Cuetos, F., & Segui, J. 2008. Phonological mediation in visual masked priming: Evidence from phonotactic repair. *Journal of Experimental Psychology:*

Variation in responses to conflicting articulatory targets in the Mandarin VN rimes

Yueh-chin Chang, Feng-fan Hsieh
National Tsing Hua University
{ycchang, ffhsieh}@mx.nthu.edu.tw

Introduction: This work is an electromagnetic articulography (EMA) study of a well-established phonotactic constraint in Mandarin (and other Sinitic languages), i.e., Rime Harmony (RH), or, a constraint dictating that the nucleus and the coda agree in backness ([1]). RH is special in that (a) restrictions of this sort are not attested in other languages since generally restrictions are not levied on vowel-coda sequences, (b) there is cross-dialectal variation in implementing RH, and more importantly, (c) unlike many other languages, it is the *velar* coda, but not the previously claimed ones such as the uvular and liquid codas ([2]), that may involve in RH, the phonotactic constraint in question (cf. [3]).

Rime Harmony: Mandarin Chinese has five phonemic vowels, /i, y, ɤ, a, u/ and two nasal codas, /n, ŋ/. Only (phonemic) vowels /i, ɤ, a/ may precede the two nasal codas. The effect of RH may be illustrated with the case of the low vowel /a/.
(i) /An/ → [an] vs. /Aŋ/ →[ɑŋ] (where /A/ means an underlyingly unspecified low vowel)
In contrast, previous impressionistic and/or acoustic studies of Taiwanese Mandarin (TM), a variety of Mandarin, report that the coda nasals undergo place neutralization when the nucleus vowel is nonlow (/i, ɤ/). The patterns are summarized in (ii). Also, the neutralizing target may be either [in] or [iŋ], depending on different regional varieties ([4], *a.o.*).

(ii) Taiwanese Mandarin	/in/	/iŋ/	/ɤŋ/	/ɤn/→[ən]
Velarization	[iŋ]	N/A	N/A	N/A
Conorization	N/A	[in]	[ən]	N/A
(iii) Standard Chinese	N/A	[iᵊŋ]	N/A	N/A

Coronization, at first blush, may be treated as a “conflict” between an advanced tongue body target for the nonlow vowels (/i, ɤ/) and a retracted target for the following tongue dorsum constriction for /ŋ/, although it is less clear how velarization may be analysed in a similar fashion. Therefore, the first goal of this study is to investigate if the variation in (ii) can be instrumentally confirmed with the help of EMA. It is also remarkable that both coronization and velarization have never been reported for the VN rimes in Standard Chinese (SC; a.k.a. Beijing Mandarin). The specific conflict is resolved by yet another possible strategy: an “excrescent schwa,” i.e., [iᵊŋ] in SC ([1], [2], *a.o.*). An articulatory study of the cross-(sub)dialectal variation helps better understand the nature of variation.

Method: Four speakers of Standard Chinese (SC) and six speakers of Taiwanese Mandarin (TM) participated in the study (aged 20-26 y.o.). They are all monolingual speakers, meaning that they don’t speak any other Chinese dialects. Kinematic data were captured using an NDI Wave with a sampling rate of 100 Hz and acoustic data were simultaneously collected during the experiments. All possible monosyllabic words containing the five monophthongs {/i/, /y/, /a/, /ɤ/, /u/} and the two nasal codas {/n/, /ŋ/} were embedded in the carrier phrase: “mà ___ ba.” ‘Scold ___ SFP’ and were produced, together with the other fillers, ten times in a randomized order. Articulatory data are processed (specifically, x=front-back; z=up-down in mm.) of the sensors attached to the tongue tip (TT), tongue blade (TB), tongue dorsum (TD), upper and lower lips (UL, LL), with the help of MView. Following [5]’s method, the trajectories of the sensors attached on the articulators in multiple dimensions over time are reconstructed, and individual variation may also be compared accordingly.

Results: From Figure 1 (Left), the data from two representative speakers show that in /pan/ and /paŋ/, TD moves upwards in SC but that’s not case in TM. In Figure 1 (Right), we see that there is an obvious back-to-front movement of TD in TM's /kɤn/; otherwise, the two varieties of Mandarin pattern alike.

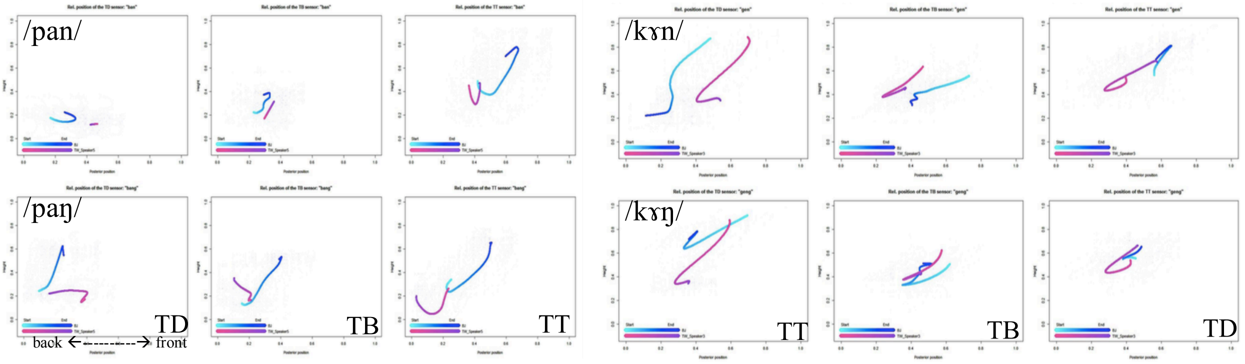


Figure 1. *Trajectoires for pan/paŋ (Left)* & *kɤn/kɤŋ (Right)*: SC in Blue and TM in Red; Light color means the beginning of the syllable, dark colored part the end of the syllable.

Regarding /iN/ rimes, cross-dialectal variation can be identified in Figure 2. Type 1 (Left) shows the so-called coronization in TM (see (ii)), whereas /pin/ and /piŋ/ in SC and TM pattern alike in Type 2 (Right). Notably, substantial back-to-front movement of TB occurs in both SC and TM in Type 2 (Right), suggesting a possible presence of the excrescent schwa (iii). Finally, velarization in (ii) is not confirmed in the present results (cf. [4]).

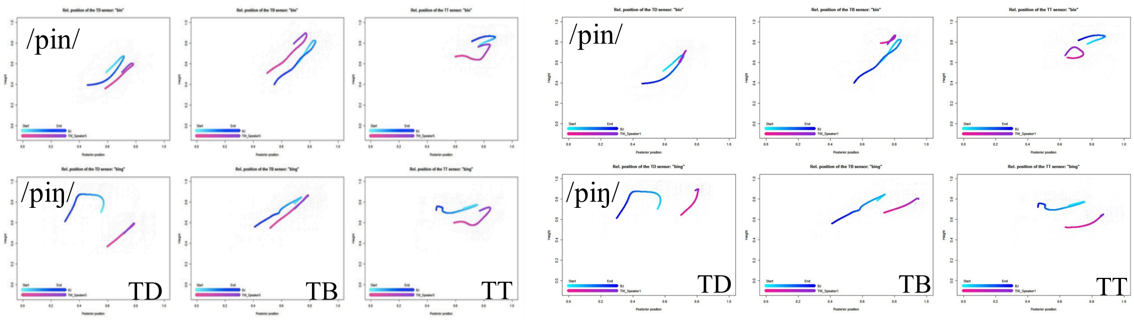


Figure 2. *Trajectoires for pin/piŋ*: Type 1*(Left)* vs. Type 2 *(Right)*: SC in blue and TM in red

Discussion : Our results indicate that the nasal codas in {/ɤn/, /in/} and {/ɤŋ/, /iŋ/} may be neutralized (Figure 1, Right and Figure 2, Left, both TM in red) and the neutralizing target is coronal. In other words, the velar may be fronted in the wake of contextual influences (see [3], *a.o.*). On the other hand, the velar may also be resistant to coarticulatory pressure, resulting in an excrescent schwa (see (iii) and Figure 2 Type 2, SC in blue). In conclusion, it is remarkable that the velar can be articulatorily "resistant" in Mandarin, resulting in a typologically distinct phonotactic constraint from other (non-East Asian) languages.

[1] Duanmu, S. 2007. *The Phonology of Standard Chinese*. OUP.
[2] Gick, B. & Wilson, I. 2006. Excrescent Schwa and Vowel Laxing: Cross-linguistic Responses to Conflicting Articulatory Targets. *Papers in laboratory phonology* VIII.
[3] Keating, P., and Lahiri, A. 1993. Fronted Velars, Palatalized Velars, and Palatals. *Phonetica* 50.2:73-10.
[4] Fon, J., Hung, J.-M., Huang, Y.-H., & Hsu, H.-J. 2011. Dialectal variations on syllable-final nasal mergers in Taiwan Mandarin. *Language and Linguistics* 12.2:273-311.
[5] Wieling, M., Tomaschek, F., Arnold, D., Tiede, M., Bröker, F., Thiele, Wood S., & Baayen, R. H. 2016. Investigating dialectal differences using articulography. *Journal of Phonetics*, 59, 122-143.

Unlearning to perceive and produce epenthetic vowels: the case of L1 Brazilian Portuguese/L2 English sequential bilinguals

Steven Alcorn (steven.alcorn@utexas.edu) and Rajka Smiljanic (rajka@austin.utexas.edu)
The University of Texas at Austin

Speech communication in second language (L2) requires both production and perception. Previous examination of the perception-production link in learning novel phonological contrasts has yielded mixed results [4, 5]. The nature of this link is equally unclear for phonotactics [2, 8]. The lack of consensus regarding the production-perception link suggests a complex relationship between the two domains in language learning. The present study examines the generalization of phonotactic constraints from L2 to L1 by looking at production and perception of vowel epenthesis in Brazilian Portuguese (BP) by learners of L2 English. More generally, it asks whether acquisition of phonotactically permissible sounds in the L2 can alter L1 representations that constrain these same sounds in the L1. In BP, the stop consonants /p, t, k, b, d, g/ are illegal in coda position and are repaired via an epenthetic /i/. This phonotactic repair is attested in both perception as an illusory vowel [7] and in production as an epenthetic vowel [3]. Importantly, no such phonotactic restriction exists in English.

Three groups of native BP speakers were tested in production and perception tasks: monolinguals living in Brazil with no or limited knowledge of English (n=15), English learners living in Brazil (n=14), and BP/English sequential bilinguals living in the United States (n=13). For production, subjects read aloud 24 sentences in Portuguese containing target words with illegal coda stops triggering epenthesis (e.g. *obter, ignorância, captar*). Target words were analyzed in Praat [1] and coded for the presence/absence of an epenthetic vowel. Figure 1 shows the rate of epenthesis for the three groups. A mixed-effects logistic regression model yielded statistically significant group differences: monolinguals produce epenthetic vowels more frequently than Brazil bilinguals ($p<.05$) and US bilinguals ($p<.001$). The difference between the L2 English speakers in Brazil and the US was not significant ($p=.074$).

The same subjects (as well as a group of L1 English control listeners, n=16) also completed a perception task in which they heard 128 BP non-words with a VC(i)CV structure produced by a female simultaneous BP/English bilingual. Half of the words contained an epenthetic /i/, 36-40 ms long, after the first consonant, while the other half had no vowel between the two consonants. Subjects completed a forced-choice identification task in which they listened to the non-words and chose the orthographic representation (e.g., <ebda> or <ebida>) that best matched the auditory stimuli (e.g., [ebda]). The results were analyzed within the signal detection framework [6]. Figure 2 shows accuracy scores (d-prime) for the four groups. A one-way ANOVA was performed with listener group as the independent variable and d-prime score as the dependent variable. The model was significant ($F(3, 54)=34.14$, $p<.0001$), and post-hoc tests revealed significant differences between all groups except the L2 English listeners in the US and Brazil.

A Spearman's rho correlation between the d-prime scores from the perception task and the rates of epenthesis from the BP word reading task was significant: $S=3693.3$, $p<.01$, $r_s=0.61$(see Figure 3). Combined, the results show that subjects with knowledge of English were able to transfer of phonotactic constraints from L2 to L1. The correlation results indicate a perception-production link suggesting susceptibility of both domains to L2-L1 influence.

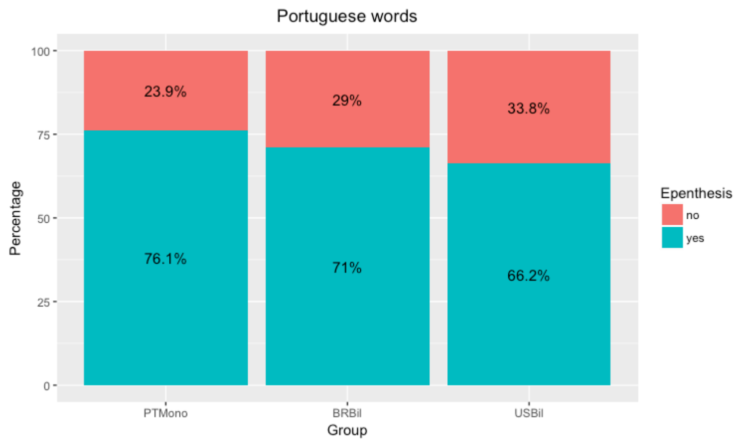


Figure 1. Rates of production of epenthesis in BP words (PTMono=BP monolinguals; BRBil=BP/English subjects in Brazil; USBil=BP/English subjects in US)

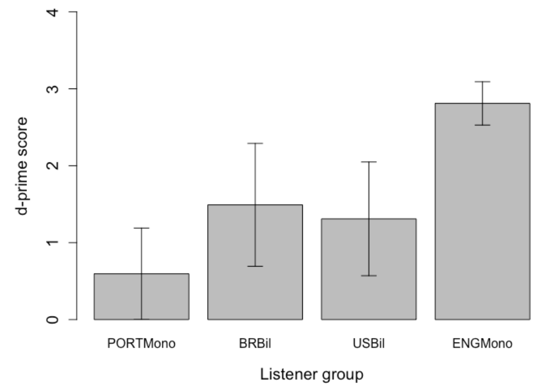


Figure 2. Average d-prime scores for listeners.

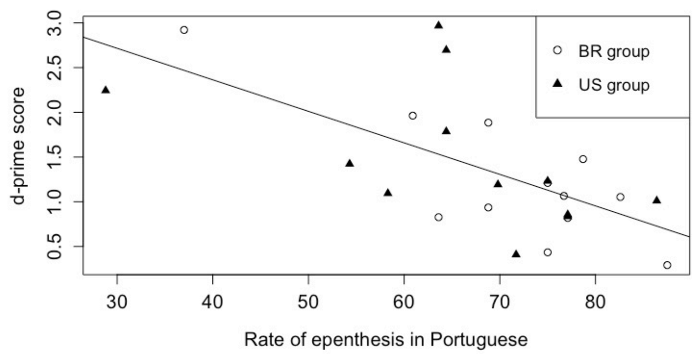


Figure 3. Scatterplot of average production (x-axis) and perception (y-axis) results for each talker/listener.

References

[1] Boersma, P. & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.36, retrieved 1 October 2017 from http://www.praat.org/.
[2] Bradlow, A.R., Pisoni, D.B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101*, 2299-2310.
[3] Câmara Jr., J.M. (1969). *Problemas de Lingüística Descritiva*. Petrópolis, Brazil: Vozes.
[4] Cardoso, W. (2011). The development of coda perception in second language phonology: a varitionist perspective. *Second Language Research, 27*(4), 433-465.
[5] Flege, J.E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.
[6] Macmillan, N.A. & Creelman, C.D. (2005). *Detection Theory: A User's Guide (2nd Edition)*. Mahwah, NJ: Lawrence Erlbaum.
[7] Parlato-Oliveira, E., Christophe, A., Hirose, Y., & Dupoux, E. (2010). Plasticity of illusory vowel perception in Brazilian-Japanese bilinguals. *Journal of the Acoustical Society of America,* 127(6), 3738-3748.
[8] Shin, D.-J. & Iverson, P. (2010). Individual differences in vowel epenthesis among Korean learners of English. *Proceedings of the ICPhS XVII*, 1822-1825.

The role of phonotactics and lexicality on the perception of intrusive vowels

Ela Portnoy and Elinor Payne
University of Oxford

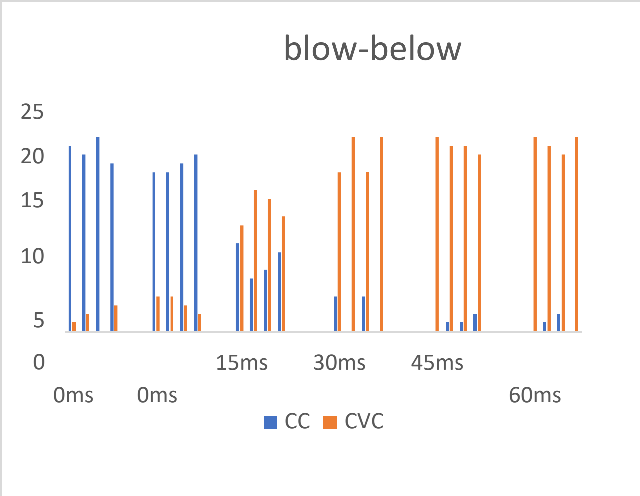
It is well-established that listeners' knowledge of phonotactics influences speech perception in a variety of ways [1,2,3]. This includes processing of consonant clusters, where it has been shown that listeners are biased towards hearing intrusive vowels in illegal clusters [4]. Less well- understood is listeners' susceptibility to hearing vocalic intrusion in *legal* clusters, i.e. where the auditory input can map onto either CC or CVC representations. It is hypothesised that the perception of such sequences is mediated by different types of linguistic information, such as lexicality and frequency, in conjunction with relevant acoustic cues. In addition, knowledge of the *phonetic implementation* of clusters, which varies cross-linguistically, may also influence perception. [5,6] show that clusters in Norwegian are characterised by an open release of C1 more frequently than their equivalent clusters in SSB English, which in turn are more characterised by reduction and overlap. This pattern is even more pronounced in early child productions, such that the dominant acquisition strategy is cluster reduction in English, and epenthesis in Norwegian. This suggests the perceptual threshold for vocalic intrusion differs cross-linguistically, even when the phonological structure is ostensibly the same.

In this paper, we investigate the influence of lexicality, lexical frequency and phonotactic legality on the perception of clusters and intrusive vowels in SSB English. The study is also a first step towards a cross-linguistic comparison with Norwegian. 22 native speakers of SSBE were asked to identity a series of auditory stimuli, in a forced choice task presented orthographically in PsychoPy. The stimuli were taken from recordings of word and non-word items with an original CVCV(C) structure, manipulated to vary V1 duration (6 different intervals: 0, 15, 30, 45 and 60 ms, plus a control 0ms token taken from a CCVC(C) production). For each stimulus, participants chose between items with CVCV(C) or CCV(C) structure, thus testing the perception or otherwise of the intrusive vowel (V1). The clusters investigated were (legal) /bl/ and /gl/, and (illegal) /dl/, and the paired items formed 13 unique lexical frames, such as "blow"/"below", "blieve"/"believe", "glow/gelow", "dlow"/"delow", "blard"/"belard". Participants heard each variant of each lexical frame 4 times, in randomised order, and their responses and reaction times were recorded.

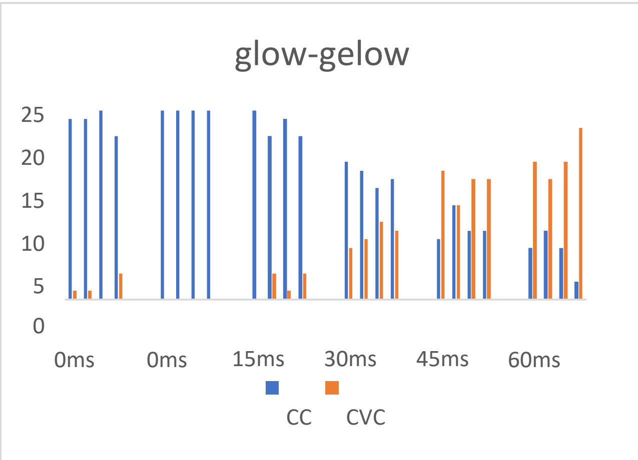
Preliminary results (see Figures 1-5) show a strong effect of both lexicality and phonotactic legality on the perceptual threshold for intrusive vowels, as well as on degree of certainty (response number and speed). Overall, the longer the intrusive vowel, the more likely it would be perceived. However, as hypothesised, the durational threshold varied according to the linguistic constraints of each lexical frame. Where *CVCV(C)* was the only real word in the frame (e.g. "blieve"/"believe"), a vocalic interval of just 15ms was needed for a clear majority response of CVCV(C) ("believe"), and RTs were quicker for CVCV(C) responses. Even where stimuli had no intrusive vowel (0ms), the CVCV(C) response was given 5-23% of the time, suggesting listeners were processing these as reduced variants of the word "believe". By comparison, where *CCV(C)* was the only real word (e.g."glow"/gelow"), intrusive vowels needed to be even longer to be perceived (at least 30ms, and 45ms for CVCV(C) ("gelow") to be the preferred response) and RTs were slower. Up to half of all stimuli with a vocalic interval of 30-60ms were heard as CCV(C) (which suggests they were mapped onto the real word "glow"), whereas stimuli with no vocalic interval very rarely yielded a response of CVCV(C) ("gelow"). Where both items were non-words, (e.g. "blard"/"belard"), a shorter interval was required for perception: with 15ms of vocalic interval, "belard" was heard with some regularity, although 30ms were needed before "belard" became the preferred response, and 45ms before being strongly preferred. However, even with 60ms there were some responses of "blard", and, analogously, at 0ms there were some responses of "belard", suggesting greater symmetry of ambiguity. When both items were non-words but one phonotactically illegal ("dlow"/"delow"), the illegal CCV(C) response ("dlow") was more strongly dispreferred overall.

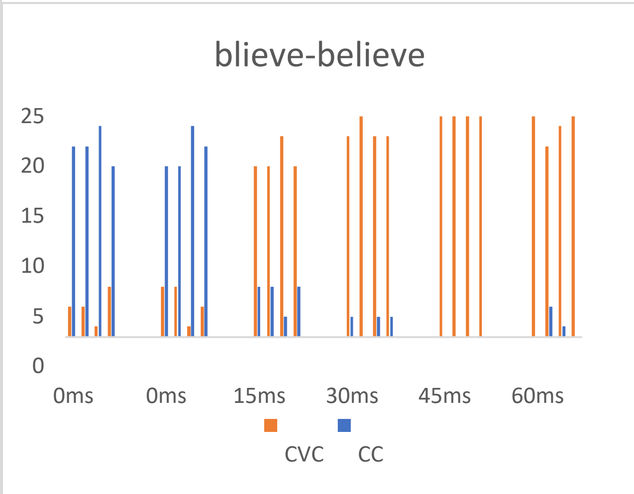
We discuss the significance of these findings for our understanding of phonological representation and in particular of phonotactic knowledge, and look ahead to possible implications for acquisition.

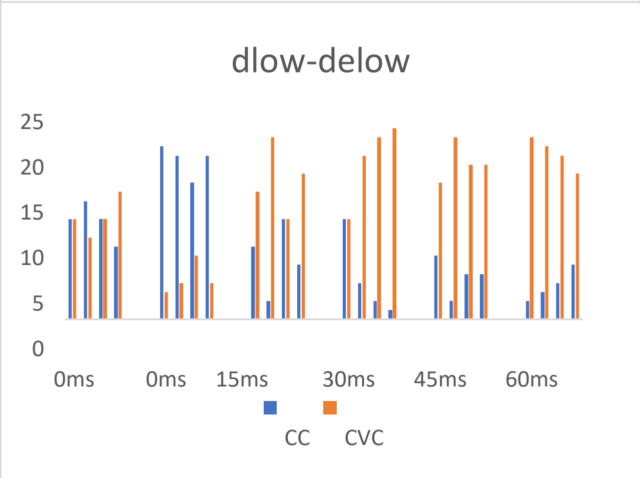
[1]Jusczyk, P. W. 1995, « Language acquisition : speech sounds and the beginnings of phonology » In J. L Miller and P. D. Eimas (Eds), *Handbook of perception and cognition, Vol 11 : Speech, language and communication*, 263-301, San Diego, Academic Press
[2] Massaro, D. M. & M. M. Cohen, 1983, "Phonological context in speech perception" in *Perception and Psychophysics*, 34:4, 338-348
[3] Davidson, L. & J. Shaw, 2012, "Sources of illusion in consonant cluste perception" in *Journal of Phonetics*, 40:2, 234-248
[4] Pitt, M. A. 1998, Phonological processes and the perception of phonotactically illegal consonant clusters, Perception and Psychophysics, 60; 6, 941-951
[5] Simonsen, H. N. Garmann, E. Payne, B. Post, E. Holm & P. Hansen, 2015, "Cross-linguistic microvariation in cluster production", oral presentation at *ISBMS*, Chania, Sept 2015
[6] Simonsen, H., N Garmann, P. Hansen, B. Post & E. Payne (*in prep*) "Vowel epenthesis in Norwegian and English clusters: a comparative study in adult and child speech"

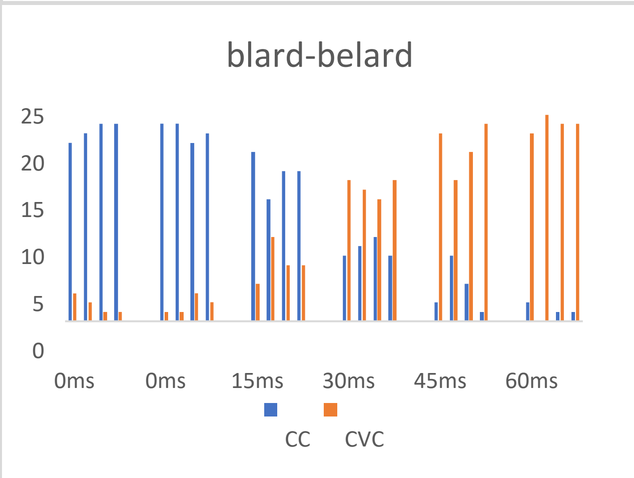


Figures 1-5: total number of responses (over 22 speakers) for repetitions 1-4, for each vocalic interval, over 5 lexical frames









Non-Māori speaking New Zealanders show surprisingly sophisticated Māori phonotactic knowledge

Yoon Mi Oh, Clay Beckner, Jen Hay, Jeanette King

New Zealand Institute of Language, Brain and Behaviour, University of Canterbury

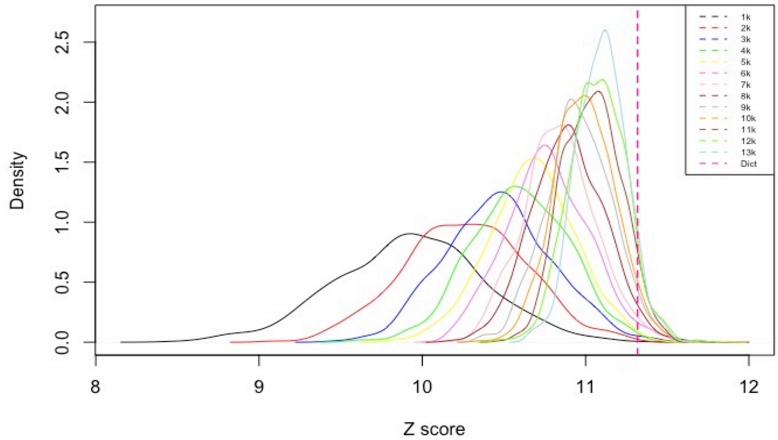
Language speakers can rate the gradient well-formedness of non-words in their language [1-2]. Such knowledge is assumed to have been acquired from statistical learning over speakers' lexicons [1-3]. Most New Zealanders (NZers) are exposed to Māori in their daily lives but do not speak Māori. They have a small lexicon comprising common loanwords and placenames. Our work has identified 121 words that most NZers can identify as Māori (although they can't define them all). We aim to understand the phonotactic knowledge that NZ-based non-Māori-speakers (NMS) & Māori-speakers (MS) have. Is it possible for non-speakers of a language to acquire statistical knowledge from only a small set of words?

Participants were asked to rate nonwords generated from a trigram model [4] for how good they would be as Māori words. We collected ratings for a total of 1760 words. Participants are 41 MS & 137 NMS. Phonotactic scores are calculated from: a Māori dictionary [5]; segmented Māori running speech data (RS) [6][7], unsegmented RS; & known words (the list of 121 words identified above, plus 55 placenames). Mixed-effects regression shows that both MS & NMS are influenced by phonotactics. The very best phonotactic predictor for both groups is the trigram model generated from the dictionary. There is no interaction between Māori-speaking status & the dictionary-derived phonotactic score. For both groups this is a much better predictor than phonotactics from known words, RS, or unsegmented RS, and both groups appear to be using these statistics equally as well. The phonotactic score derived from known words exhibits a significant difference between MS & NMS, such that the former appears to be less influenced by this than the latter. One possible reason that this is an inferior predictor to the dictionary is that the smaller size of the known-word set makes it a less robust training set. To further assess this question, we conduct Monte Carlo simulations using 1k random samples of 150 words from [5]. Our results again show that the known words are a better predictor than a random selection from the dictionary for NMS, but not MS, but the dictionary is the best predictor for both groups. Perhaps this is because a relatively small lexicon actually approximates the statistics of the whole dictionary.

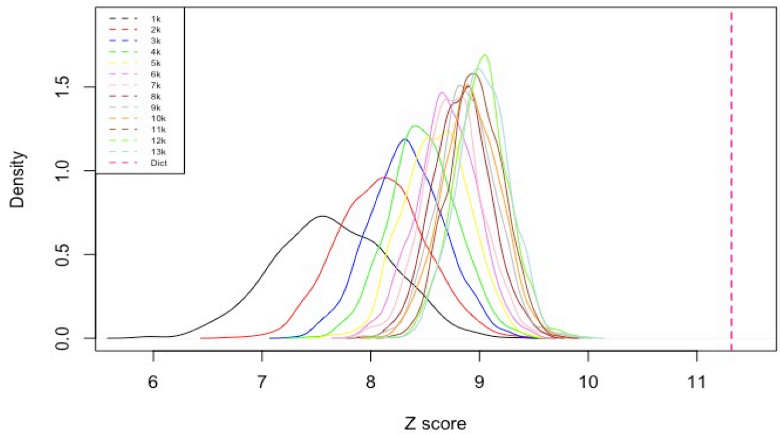
How big a lexicon would you need to have, to generate a similar trigram model to the dictionary-derived one? We tested Monte Carlo simulations with 1k random samples from [5] while varying the size of samples from 1k to 13k (c.f. Fig 1), using mixed-effects regressions to predict the ratings of NMS. As the vocabulary size increases, the z-scores of random samples gradually converge toward a full dictionary. This shows that NMS' actual phonotactic ratings of nonwords are successively better approximated by increasingly large samples of the Māori dictionary—with the very best predictions still provided by the full dictionary. This simulation appears to show that NMS' Māori phonotactic knowledge is best explained if we assume they have access to the full Māori lexicon. A similar set of simulations was done with frequency-rated random samples, with words selected into the individual samples in proportion to their lexical frequency. Phonotactics derived from these small lexicons perform still worse (c.f. Fig 2). Thus, we succeeded in our goal of demonstrating that speakers with a very small lexicon can have quite sophisticated phonotactic knowledge of a language. However we are left with the paradox that the knowledge appears to be too good to be true. To explain this result, we are exploring the possibility that the limited vocabulary of NMS provides enough initial phonotactics to allow for segmentation of ambient running speech, and this segmentation then leads to statistics derived from a much larger word-base than they appear to know. Regardless of the explanation, we have certainly found that non-speakers of a language can generate sophisticated phonotactic knowledge, and that phonotactics need not arise as a generalization over a large and established lexicon.

FIGURES

(1) Density plots from Monte Carlo simulation with different sizes of vocabularies (dictionary)



(2) Density plots from Monte Carlo simulation with different sizes of vocabularies (frequency-weighted dictionary)



Figures 1 & 2: Each colour shows a distribution of 1000 z-scores from mixed effects regression models, predicting NMS's ratings from a phonotactic score. The scores are generated by random samples of the dictionary (Fig 1) or by frequency-weighted random samples of the dictionary (Fig 2) to simulate vocabularies of different sizes.

[1] Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: effects of segment probability & length on the processing of nonwords. *Journal of Memory & Language*, *42*(4), 481–496. https://doi.org/10.1006/jmla.1999.2692.

[2] Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics & neighborhood activation in spoken word recognition. *Journal of Memory & Language*, *40*(3), 374–408.

[3] Adriaans, F., & Kager, R. (2010). Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory & Language*, *62*(3), 311-331.

[4] Needle, J., Pierrehumbert, J. & Hay, J. (2014). Phonotactic probability & wordlikeness: a flexible pseudoword generator with triphones, Madison, WI, USA: *19th Mid-Continental Phonetics & Phonology Conference,* 12-14 Sep 2014.

[5] Moorfield, J. C., (2005). *Te aka: Māori-English, English-Māori dictionary & index,* Auckland, N.Z: Pearson Longman.

[6] King, J., Maclagan, M., Harlow, R., Keegan, P. & Watson, C. (2011). The MAONZE project: changing uses of an indigenous language database. *Corpus Linguistics & Linguistic Theory, 7*(1), 37-57.

[7] Boyce, M. (2006). *A corpus of modern spoken Māori.* PhD Thesis, VUW.

The adaptation of native clusters with non-native phonetic patterns is task-dependent

Harim Kwon[1,2] Ioana Chitoran[2]
[1] *George Mason University* [2] *Université Paris Diderot*
[1] hkwon20@gmu.edu [2] ioana.chitoran@univ-paris-diderot.fr

It has been repeatedly shown that non-native consonant clusters are modified to conform with native language phonotactics both in perception and production (e.g., [1, 4, 5, 6]). This study asks (1) whether licit onset clusters with non-native phonetic patterns are also modified, or adapted, to match the native patterns, and (2) whether different tasks can induce a different degree of adaptation. We examine how native speakers of Georgian (a rich onset cluster system) spontaneously imitate word-initial clusters produced by a French talker in two conditions: Word-form Shadowing (**WS**) and Sentence Completion (**SC**). Georgian has different phonetic implementation of onset clusters from French: Georgian has longer inter-consonant timing lag [2, 7], which often results in a transitional schwa [3]. In addition, Georgian has an initial prominence for CVCV forms while French has a final prominence. Based on these differences, we predict that, if non-native phonetic patterns are adapted to native patterns, Georgian speakers will (1) produce transitional vowels when imitating French CCV, and (2) imitate French CV_1CV_2 sequences with "illusory clusters" especially when V_1 is similar to the transitional vowel produced in Georgian native clusters.

Participants were 25 native speakers of Georgian living in Tbilisi, Georgia. They were randomly assigned into two experimental conditions.

Stimuli: A French native talker produced 32 $C_1V_1C_2V_2$ pseudo-words with 8 different C_1C_2 combinations (ps, pt, sk, sp, bl, gl, pl, kl). V_1 alternated among /a/, /u/, /ø/, and no vowel, and V_2 was always /a/ (e.g., /pasá/, /pusá/, /pøsá/, and /psá/). The "no vowel" stimuli were essentially monosyllabic /$C_1C_2V_2$/. Acoustic analysis of the stimuli revealed that (1) French "no vowel" stimuli never had a transitional vowel, and (2) French /ø/ was acoustically similar to schwa in its formants (mean F1=413Hz, F2=1605Hz, F3=2584Hz).

Task: In the **WS** condition, 14 participants (1) saw CCV/CVCV sequences in Georgian script and read them aloud (baseline), and (2) heard and shadowed (immediately repeated what they heard without being told to "imitate") the French auditory stimuli. In the **SC** condition, 11 participants were asked to produce the target CCV/CVCV sequences embedded in a Georgian carrier phrase "ვეება ___ ფუჩია" /veeba ___ puʧʼia/. The participants (1) saw the carrier phrase with the target sequences in Georgian script and read them aloud (baseline), and (2) heard the French stimuli while seeing the carrier phrase with an empty slot, and produced the carrier phrase completed with the heard target sequence (test).

Results: Preliminary results (22 speakers analyzed so far [**WS**=14, **SC** =8]) suggest that segmentally native onset clusters with non-native phonetic patterns were indeed adapted, but to different degrees in different tasks. As predicted, the participants' test productions reflected modifications of French stimuli towards their native (baseline) productions, such as transitional vowels that do not exist in the auditory target, or "illusory clusters" when imitating French CVCV sequences. Moreover, these modifications were more frequent in **SC** than in **WS**. The transitional vowels were less frequent in both test conditions than in their baselines, but this decrease was significantly smaller in **SC** than in **WS** (χ^2 =5.1, p = .02, **Fig.1**). "Illusory clusters" were also more frequent in **SC** than in **WS** (χ^2 =11.9, p < .01), occurring almost exclusively when V_1 was /ø/ (e.g., /pøta/ imitated as /pta/) in both conditions (**Fig.2**).

Taken together, we claim that the effects of native language on adaptation of word-initial consonant clusters are not limited to their segmental composition, but also involve their phonetic implementation, such as timing lag and the occurrence of transitional vowels. The current findings also suggest that producing sentences in one's native language induces more rigorous modifications from the auditory targets than producing words in isolation (c.f., [4]).

(1)

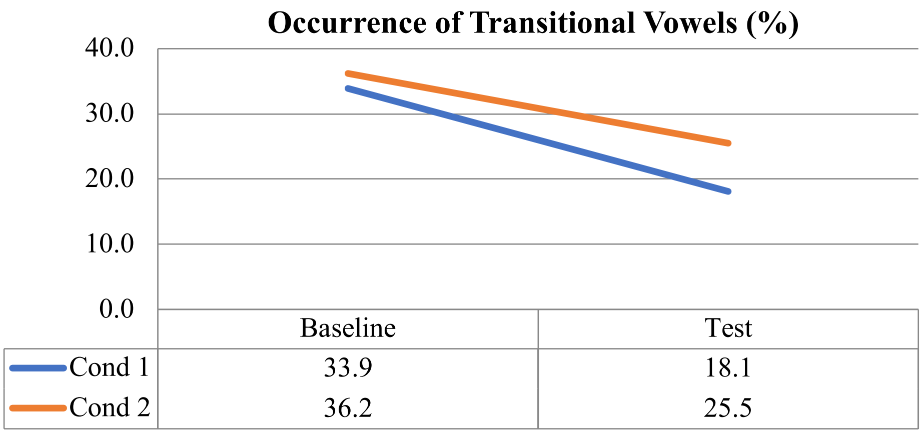


Figure 1. *Occurrence of transitional vowels produced within CCV sequences*

(2)

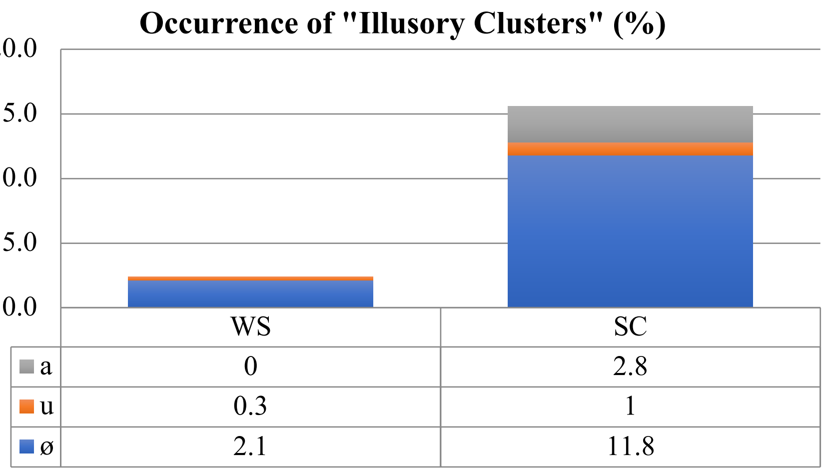


Figure 2. *V_1 in French stimuli when "illusory clusters" were produced*

[1] Berent, I., Steriade, D., Lennertz, T. & Vaknin, V. 2007. What we know about what we have never heard: Evidence from perceptual illusions. Cognition 104, 591-630.

[2] Chitoran, I., Goldstein, L. & Byrd, D. 2002. Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven, T. Rietveld & N. Warner (eds.) Laboratory Phonology 7. Berlin: Mouton de Gruyter. 419-447.

[3] Chitoran, I. & Kwon, H. 2016. Timing lag matters in the perception of Georgian stop sequences by native speakers. Presented at LabPhon15. Ithaca, NY.

[4] Davidson, L. 2003. The atoms of phonological representation: Gestures, coordination and perceptual features in consonant cluster phonotactics. Unpublished PhD dissertation, Johns Hopkins University.

[5] Davidson, L. 2006. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. Journal of Phonetics 34, 104-137.

[6] Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C. & Mehler, J. 1999. Epenthetic vowels in Japanese: A perceptual illusion? Journal of Experimental Psychology: Human Perception and Performance 25, 1568-1578.

[7] Hoole, P., Bombien, L., Kühnert, B. & Mooshamer, C. 2009. Intrinsic and prosodic effects on articulatory coordination in initial consonant clusters. In G. Fant, H. Fujisaki & J. Shen (eds.) Frontiers in Phonetics and Speech Science. The Commercial Press. 275-287.

Machine learning shows that clusters representations have a temporal component

[1]Tom Lentz, [2]Marianne Pouplier, [3]Phil Hoole
[1]University of Amsterdam; [2,3]IPS Munich, Ludwig Maximilians Universität
[1]T.O.Lentz@uva.nl , [2]pouplier@phonetik.uni-muenchen.de, [3]hoole@phonetik.uni-muenchen.de

Languages differ not only in their inventory of consonant clusters [1,2], but also in the coarticulation, or timing, of such clusters [3]. Thus the same segmental sequence can be articulated quite differently across languages. There are different proposals on the underlying control structures that give rise to the observed articulatory differences [4], but as hypotheses about timing patterns exist in a high-dimensional space the best formalisation might still be undetected. The high dimensionality arises because coordination between consonants might be relative to onsets, offsets, and/or plateaus, yielding a myriad of logically possible coarticulation patterns. Without knowing which underlying relation best describes the temporal component of cluster representations, it is hard to know if temporal representations truly differ between languages and if speakers can adapt their timing to another language. However, deep learning techniques allow to separate patterns with minimal a priori assumptions about them. The present study explores Support Vector Machines (SVM), in which two categories are separated by an (*n-1*)-dimensional hyperplane fitted computationally through two sets of *n*-dimensional data points representing each category [5]. The separation can be non-linear (with the *kernel trick)* Although in principle not hidden, SVM fits do not offer insight into the way the categories are separated, only into the practical separability. However, if two datasets can be separated on given dimensions, it is theoretically possible to find a representation of the categories based on these dimensions.

We illustrate the usefulness of SVMs with data from a study comparing CCV onset clusters recorded from ten Georgian and eight German native speakers. The languages have been claimed to differ in their consonant cluster timing pattern, with Georgian having lower consonant overlap [6] than what has independently been reported for German [7]. In our study, each speaker shadowed cluster productions from two auditorily presented model speakers, one German, one Georgian. In the native condition, participants heard the model (and hence coarticulatory pattern) corresponding to their native language whereas in the non- native condition, participants imitated the model from the other language. All recorded clusters were phonotactically legal in both languages, but differed in their temporal overlap between the consonants depending on whether they were spoken by the German or Georgian model. Articulatory movement data (EMA) were recorded. For each consonant and vowel six articulatory landmarks were identified. Time points were registered relative to the velocity peak of the first consonant, yielding 17 other measurements per production for 602 cluster produced by German speakers (160 native productions, 442 non-native, i.e. imitations of Georgian) and 767 of Georgian speakers (230 native, (537 non-native). The native productions in the data were divided randomly in a training (95%) and a test (5%) set. An SVM was fitted to the training set, separating German and Georgian. It then classified the test set, as well as the non-native productions. The procedure was repeated 20000 times (1000 x 20-fold crossvalidation). Figure 1 shows the performance on the test set is good (d' = 3.8, 95% CI: 3.84-3.90). Figure 2 shows how German's non-native productions fooled the SVM: most were classified as Georgian (p= 0.54; for all reported differences $p<0.001$***). Yet Georgians' non-native productions were mostly *not* classified as Germans when imitating the German model (p = 0.26), meaning they did not reach the German coarticulation pattern.

Results confirm that Georgian and German differ substantially on some temporal component of CCV articulation. Some speakers (German participants) can imitate the timing pattern of another language, indicating that this part of the representation is not completely opaque to them. We will discuss possible reasons for the asymmetric result between languages.

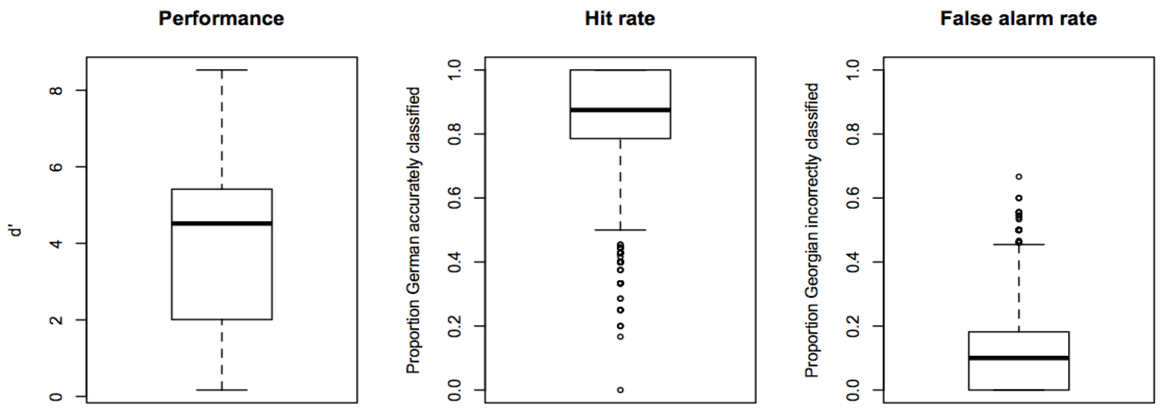


Figure 1. *Model performance on test set over 20000 runs; box indicates first to third quartile, whiskers extend to 1.5 interquartile range*.

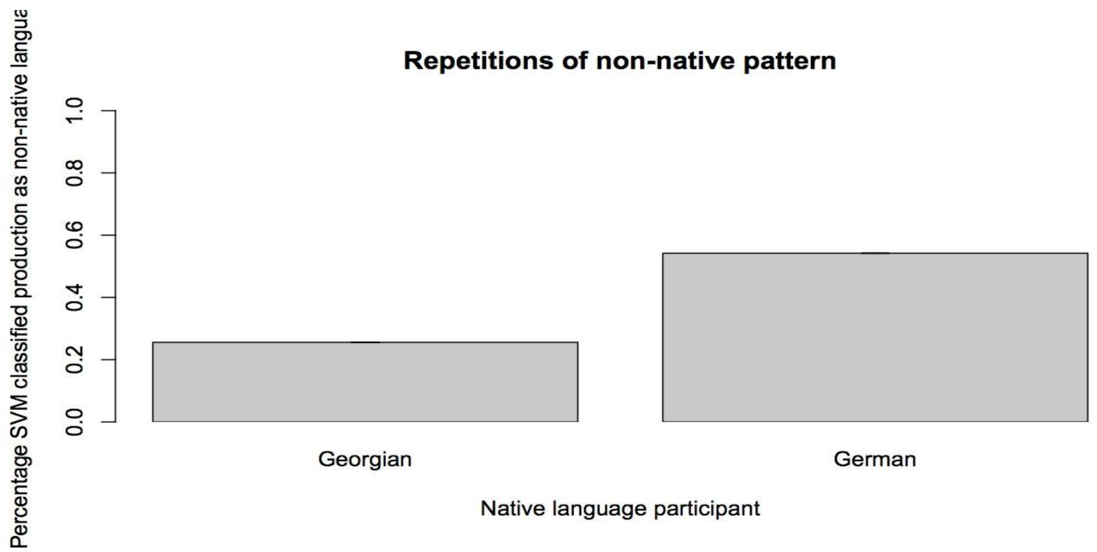


Figure 2. *Model's classification of speakers imitation of the other language as that language, per participant group, over 20000 models each fit to 95% of the native productions. Georgians repeating German are classified as German less than Germans repeating Georgian are classified as Georgian. Error bars (hardly visible) indicate s.d.*.

[1] Trubetzkoy, N. S. 1939. Travaux du Cercle Linguistique de Prague 7: Grundzüge der Phonologie. Prague: Jednota Cěskoslovenských Matematiků a Fysiků.
[2] Chomsky, N. & Halle, M. 1965. Some controversial questions in phonological theory. Journal of Linguistics, 1, 97–138.
[3] Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory*, 20, 269—337.
[4] Browman, C. P. and Goldstein, L. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica*, 45, 140–155.
[5] Bennet, K. T. and Campbell, C. 2000. Support Vector Machines: Hype or hallelujah? *SIGKDD Explorations*, 2(2), 1–13.
[6] Chitoran, I., Goldstein, L., and Byrd, D. 2002. Gestural overlap and recoverability: Articulatory evidence from Georgian. *Laboratory Phonology*, 7, 419–447.
[7] Bombien, L., Mooshammer, C., & Hoole, P. (2013). Articulatory coordination in word-initial clusters of German. Journal of Phonetics, 41(6), 546-561

Beyond phonology: The role of distal suprasegmental rate and rhythm in phonotactic parsing of speech

Laura Dilley
Dept. of Communicative Sciences and Disorders, Michigan State University
idilley@msu.edu

A robust finding is that the parsing of a given chunk of the acoustic speech signal can be heavily influenced by suprasegmental characteristics around that speech chunk. That is, it is now well-replicated that the rate and rhythm of distal, i.e., temporally nonadjacent, speech context can induce dramatic perceptual readjustments of how many syllabic units are perceived to be present in a chunk of acoustic signal, as well as in the phonemic composition and phonotactic structuring and sequencing of those units. For example, previous results show that distal context speech rate influences whether a phrase with an embedded reduced function word, e.g. "or" in the phrase "leisure or time", is heard as containing a function word or not, implying that the same acoustic chunk of speech is parsed in different ways under different distal speech rates. These distal speech rate effects on parsing the acoustic speech signal have now been found in English, Russian, and Mandarin Chinese. Here we test the hypothesis that linguistic competency in perception and production entails knowledge of statistical dependencies between the timing of syllabic units and their rates of occurrence, with implications for phonotactic parsing of speech into units – phonemes, syllables, and words. We report on three experiments that tested whether distal speech rate affects perceived phonotactic structuring of not just function words, but also of other kinds of morpho-phonological contexts. In Experiment 1, we examined whether distal speech rate influenced lexical perception for content words differing in number of syllables, e.g., form vs. forum. In Experiments 2 and 3, we used separate tasks to examine whether distal speech rate influenced perception of a reduced vowel, causing reorganization into different lexical and phonotactic units (e.g., cease, see us). Results showed that that distal speech rate significantly influenced perception of lexical content in both experiments. These findings demonstrate that distal rate substantially influences how listeners perceive structure – including phonotactic sequencing and phonemic composition – for a wide variety of phonological contexts and lexical materials. Taken together with corpus evidence supporting statistical dependencies between distal context speech rate and proximal syllable duration, these findings support the view that language competency entails inferences about the phonological composition (e.g., number of syllabic and phonemic units) for sonorous speech material, as well as the phonotactic sequencing of those units. These inferential processes about phonotactic sequencing reveal themselves especially under experimental conditions when evidence of acoustic "landmarks" (i.e., spectral discontinuities) within proximal speech are minimal. It is suggested that data explanation and/or predictive coding approaches to language perception and production provide a means of accounting for these effects of distal speech rate and rhythmic context on phonotatic parsing of speech material.

Syllable boundary displacement in Limbu verb-stem alternations

Boyd Michailovsky
Lacito, CNRS, and LabEx Empirical Foundations of Linguistics, France

Phonotactics is critical to any coherent phonology of Limbu, a Tibeto-Burman language of the Kiranti group spoken in the hills of Eastern Nepal and neighboring India. In fact, an important morphophonemic alternation in Limbu is best described as in part consisting in the movement of a syllable boundary. (Examples below are in the Mewa Khola dialect, but the principle holds for all dialects.)
Limbu syllable-initials include two series of stops, unaspirated and aspirated. There is no phonological opposition of voice; stop voicing ([p] vs [b], [pʰ] vs [bʱ] (etc.)) is determined by context. I transcribe voicing as it is normally pronounced.
The inventory of syllable-finals is basically: p, t, k, m, n, ŋ, zero. The stop finals are pronounced unvoiced and unreleased, with simultaneous glottal closure. (Stop finals are not found between vowels except before a morpheme boundary.)
Syllables are (Ci)V(Cf), with no clusters. Words are made up of one or more syllables.
There is a general phonological rule that syllable-initial stops are voiced after a vowel or nasal final (even across a word boundary): paːn ‘speech’, kubaːn ‘his speech’
A further rule is that word-finals are geminated before a vowel-initial suffix, Exx: kumeːt ‘his wife’, kumeːtti? ‘his wife-Q’, kumeːttaŋ ‘his wife too’; kɛbeːk ‘you go’, ‘kɛbeːkki’ ‘will you go?’ (stem: peːk).
Verbs have complex affixal morphology. The root has the (morphological) form (C1)V(C2)(C3), where C3 (the “augment”) is either t or s. (In word families, these are formatives with applicative (t) or causative (s) semantics.) Each lexical verb has two alternating stems, which can be called ‘non-past’ and ‘past’. Non-past stems have the canonical form (Ci)V(Cf), and can stand alone as phonological words (haːp ‘he weeps’). Past stems have the form (Ci)V(Cf)Ci-; the obligatory Ci at the end corresponds to C2 or C3 of the root (or is an epenthetic j with a CV root) and must be followed by a vowel-initial suffix string. Typical stems and forms (hyphens separate morphemes; dots (past stem only) separate syllables):

root	gloss	Non-past stem exx.		past stem exx.	
juŋ	‘stay’	mɛn-juŋŋ-ɛ-aŋ	‘not having stayed’	ju.ŋ-ɛ	‘he stayed’
nuːks	‘return’	mɛn-nuːŋŋ-ɛ-aŋ	‘not having returned’	nuːk.s-ɛ	‘he returned’
haːp	‘weep’	mɛn-haːpp-ɛ-aŋ	‘not having wept’	haː.b-ɛ	‘he wept’
hipt	‘strike’	mɛn-hipp-ɛ-aŋ	‘not having struck it’	hip.t-ɛ	‘he struck it’
tsok	‘do’	mɛn-dzoːkk-ɛ	‘don’t do it’	tsoː.g-ɛ	‘do it!’
tɔ	‘dig’	mɛn-dɔ-ɛ	‘not having dug’	tɔ.j-uŋ	‘I dug it’

Weidert & Subba (1985) recorded the Panchthar forms correctly but could not find a phonological principle that predicted whether the apparent stem final in a given form would be geminated (e.g. Panchthar haːppa ‘I weep’) or voiced (Panchthar haːba ‘he wept’). This was mainly a failure to take phonotactics into account: if the stems are analysed according to the stem canons stated above, which hold for all regular verbs, it falls out that Cf (i.e. at the end of non-past stems) are geminated intervocalically and Ci stops (at the end of CV.C past stems) are voiced, as in the rest of Limbu phonology. The stem alternations (not all types are shown here) are applicable only to verb stems, and are not part of the phonology proper. But the stem alternation of CVC roots consists essentially in the displacement of the syllable boundary. In the proposed analysis, it is the syllable structure that determines the realization of the segments.

Probabilistic phonotactics in visual word recognition within and across morphological boundaries

Chiara Celata*, Giulia Bracco°
Scuola Normale Superiore di Pisa, °Università di Salerno

Transition probabilities (TPs) refer to the conditional probability with which a segment σ occurs in a corpus, given *N* preceding segment(s). TPs are intuitively similar to a frequency measure but produce an opposite effect on word recognition. High TPs imply high neighborhood density and therefore inhibitory effects on word or syllable recognition. TPs have been shown to influence infants' discrimination of native language sounds and words, children's and adults' spoken word recognition, wordlikeness ratings, speech segmentation and speech production (e.g. [1], [2], [3], [4], [5], [6]).

This study investigates the role of TPs in the processing of consonant clusters (as opposed to consonant+vowel sequences) and of complex words (as opposed to monomorphemic words) in two experiments of visual word recognition. Consonant clusters are challenging phonological objects not only for phonetic complexity but also for cognitive processing [7] and are known to convey information on the morphological structure of words ([8], [9] [10]).

Two fragment priming lexical decision tasks, in which an Italian target word (e.g. *borsa* 'bag') is primed by its initial fragment (*bor*), are performed by 124 native speakers. In both experiments, based on a reference corpus of written Italian [11], we calculate TP and token frequency values for each biphone corresponding to the last segment of the fragment prime and the following segment in the target (i.e. ***rs*** in *bo**rs**a*). TP and frequency values are equally calculated for the sequence corresponding the fragment prime and the following segment (i.e. ***bors*** in ***bors**a*). In the first experiment (E1), biphones are CV or CC sequences (e.g. *bir-bi**ri**llo* 'skittle' vs. *bor-bo**rs**a* 'bag') and the aim is to verify whether biphone's and sequence's TPs play a different role in consonant clusters as opposed to consonant+vowel sequences. The second experiment (E2) focuses on consonant clusters and compares monomorphemic words (e.g. *bis-bi**st**ecca* 'beef-steak') with prefixed words in which the biphone crosses a morphemic boundary (e.g. *bis-bi**sn**onna* 'great-grandmother', 'bis' being an Italian prefix). Linear mixed models [12] are run with log-transformed reaction times as the dependent variable, Target Frequency, Prime Frequency, Biphone Frequency, Biphone TP, Sequence Frequency and Sequence TP as independent factors with fixed effects, Word and Subject as factors with random effects.

The results of E1 (Figure 1) show that latencies in the recognition of CC and CV targets are predicted by different factors. Target Frequency has a strong role in both CC and CV models; however, the models differ for the role of sublexical factors. In particular, there is an inhibitory effect of Sequence TP on target recognition in the CC model, suggesting that probabilistic information is processed in the recognition of targets containing consonant clusters and not of targets containing CV sequences. The results of E2 additionally show that probabilistic and frequency-related information influence the processing of prefixed and monomorphemic words differently. Sequence TP has an inhibitory effect on the recognition of monomorphemic targets, consistently with the CC results in E1, but an opposite facilitatory effect is found in the recognition of prefixed words. Moreover, both Target Frequency and Prime Frequency have a facilitatory effect on prefixed words, whereas Prime Frequency has an inhibitory effect on monomorphemic words, suggesting a competition effect for the latter but not for the former.

The results of this study enlarge the scope of current investigation on probabilistic phonotactics, showing, for an understudied language, the role of TPs in the processing of consonant clusters and of morphologically complex words. Implications for the phonotactics-morphology interface and for theories of visual word recognition are also discussed.

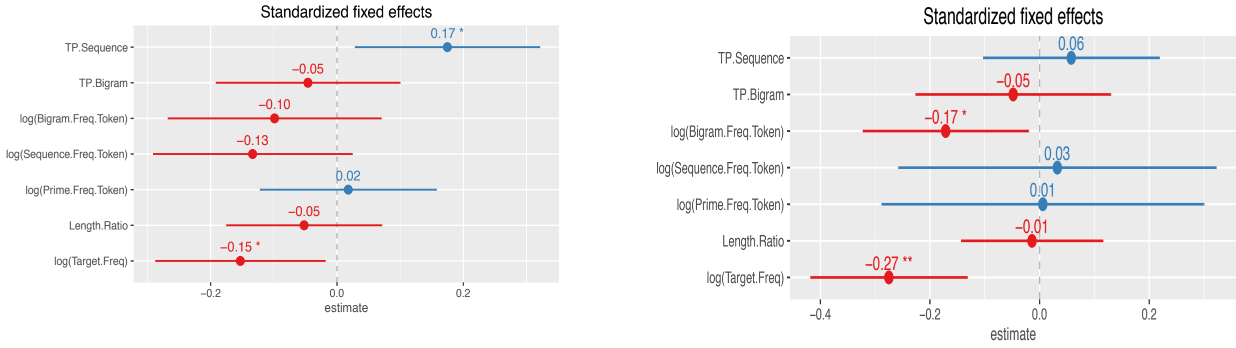


Figure 1. *Standardized fixed effects for the CC (left) and CV (right) models in E1. Stars (*) represent the significance levels : '***' = 0.001, '**' = 0.01 , '*' = 0.05.*

[1] Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630-645.

[2] MacRoy-Higgins, M., Shafer, V.L., Schwartz, R.G. & Marton, K. (2014) The influence of phonotactic probability on word recognition in toddlers. *Child Language Teaching and Therapy* 30(1), 117-130 .

[3] Vitevitch, M.S., Luce, P.A., Charles-Luce, J. & Kemmerer, D. (1997) Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech* 40, 47-62.

[4] Mattys, S.L. & Jusczyk, P.W. 2001. Do infants segment words or recurring contiguous Patterns? *Journal of Experimental Psyhcology: Human Perception and Performance* 27(3), 644-655.

[5] Gustafson, E. & Goldrick, M. 2017. The role of linguistic experience in the processing of probabilistic information in production. *Language, Cognition and Neuroscience*, 33(2), 211-226.

[6] Vitevitch, M. S., & Luce, P. A. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408.

[7] Hoole, P., Bombien, L., Pouplier, M., Mooshammer, C. & Kühnert, B. 2012. *Consonant clusters and structural complexity*. Mouton de Gruyter.

[8] Dressler, W.U. & Dziubalska-Kolaczyk K. 2006. Proposing morphonotactics. *Italian Journal of Linguistics* 18, 249-266.

[9] Calderone, B., Celata, C., Korecky-Kröll, K. & Dressler, W.U. 2014. A computational approach to morphonotactics: evidence from German. *Language Sciences* 46 (A), 59-70.

[10] Celata, C., Korecky-Kröll, K., Ricci, I. & Dressler, W.U. 2015. Phonotactic processing and morpheme boundaries: word- final /Cst/ clusters in German. *Italian Journal of Linguistics* 27(1), 85-110.

[11] Bertinetto, P.M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C. & Thornton, A.M. 2005. *Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS)*. http://linguistica.sns.it/CoLFIS/Home.htm

[12] Baayen, R.H., Davidson, D.J. & Bates, D.S. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

Grammaticality and Lexical Statistics in Chinese Unnatural Phonotactics

Donald Gong

University College London (UCL) London, UK
ucjudgo@ucl.ac.uk

Speakers possess phonotactic knowledge about the acceptability of non-words, yet the source of this knowledge is unclear. One possibility is that a non-word is judged to be unacceptable because it violates the phonotactic grammar of this language. For instance, syllables in Standard Chinese take the form of CGVX (G=glide, X=vowel length, glide or nasal). To account for the syllable phonotactics of Chinese, four OCP-based phonotactic constraints can be proposed as part of such a phonotactic grammar, under the assumption that a natural phonotactic constraint is either 1) phonetically grounded, or 2) typologically well-attested (Hayes and White, 2013):

(1) Phonotactic Constraints in Chinese	Example
*HH: The feature [+high] cannot occur in sequence.	*[l**ui**] *[t**yu**]
*[Cor]_[Cor]: [Cor] cannot occur in both G and X.	*[**j**a**i**] *[p**jei**]
*[Lab]_[Lab]: [Lab] cannot occur in both G and X.	*[**w**o**u**] *[nwa**u**]
C and G must have different articulators	*[**tʂj**an] *[**pw**aŋ]

Another possible account for the acceptability judgments is based on how similar the non-word is to all real words in the lexicon. Multiple models have been proposed to capture this analogical effect, and we focus on two of them: the Neighbourhood Density model (Bailey and Hahn, 2001) and Hayes & Wilson's Phonotactic Learner (Hayes and Wilson, 2008). Neighbourhood Density counts the number of words generated by substituting, deleting, or adding a single phoneme together with their summed frequency. For example, the form *lat* has abundant lexical neighbours in English (e.g. *cat, lap*), while *zev* has a sparse neighbourhood density. Phonotactic Learner produces a set of feature-based constraints given a feature matrix and a lexicon for training. The learner attempts to identify the constraint set and a set of constraint weights that maximise the probability of the input forms. We could then apply this learned grammar to evaluate the grammaticality of non-words by assigning penalty scores.

Linear logistic regression analyses were applied to the data of a phonological acceptability judgement mega study run on 110 Mandarin native speakers (Myers and Tsay, 2015). We used Neighbourhood Density, penalty scores generated by the Phonotactic Learner, and whether the phonotactic constraints in (1) are violated as independent variables to predict speaker's reaction time on the lexical decision task, with Neighbourhood Density and Phonotactic Learner representing lexical statistics, and constraints in (1) representing grammaticality. Results show that each parameter plays an independent role, suggesting that even though lexical statistics and grammaticality overlap substantially, each still independently contributes to speaker's reactions (Table 1). The results suggest that the extreme lexicalist view, which attributes all phonotactic patterns to frequency statistics (Hay, Pierrehumbert and Beckman, 2003) is too strong.

Non-words that violate the constraints in (1) are labelled as systematic gaps, while other missing syllables are labelled accidental gaps. However, some of the accidental gaps are not so 'accidental' as expected. We noticed a specific phonotactic constraint that bans the cooccurrence of a labial fricative with a following coronal glide (*[fj]), and incorporated it into the statistic model. Despite the constraint's phonetic unnaturalness, the reaction time results suggest that speakers reject *[fj] gaps more quickly than other accidental gaps, as if they were systematic gaps (Figure 1). Therefore, the relevance of this constraint in Chinese indicates that, unlike what has been proposed by Becker et al. (2011), unnatural phonotactics can be learned by speakers and be part of the phonotactic knowledge. The possibility that *[fj] is a natural constraint or that it is a result of the phonemic analysis of Standard Chinese adopted here, however, will be discussed

	β	SE(β)	z	p
(Intercept)	-0.7212	0.0305	-23.730	
penalty	-0.0052	0.0020	-2.570	.0102*
neighbourhood density	0.0101	0.0028	3.578	.0003*
being a systematic gap	-0.0448	0.0212	-2.108	.0350*
penalty : *neighbourhood density*	-0.0007	0.0002	-3.093	.0020*

Table 1 Results of linear logistic regression on response

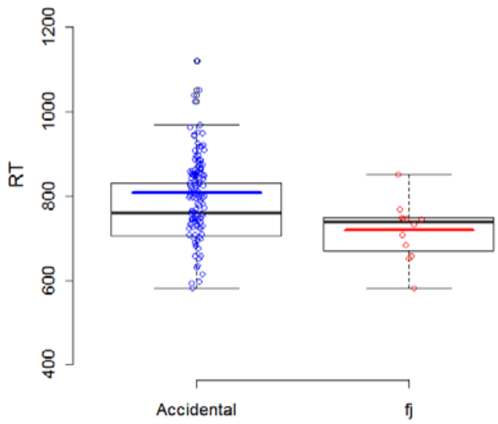


Figure 1 Reaction time distribution of accidental vs. *[fj] violating gaps

References

Bailey, T. M. and Hahn, U. (2001) 'Determinants of Wordlikeness: Phonotactics or Lexical Neighborhoods?', *Journal of Memory & Language*, 44(4), p. 568.

Becker, M., Ketrez, N. and Nevins, A. (2011) 'the Surfeit of the Stimulus: Analytic Biases Filter Lexical Statistics in Turkish Laryngeal Alternations', *Language*, 87(1), pp. 84–125.

Hay, J., Pierrehumbert, J. and Beckman, M. E. (2003) 'Speech perception, well-formedness and the statistics of the lexicon', in Local, J., Ogden, R., and Temple, R. (eds). Cambridge: Cambridge University Press, pp. 58–87.

Hayes, B. and White, J. (2013) 'Phonological naturalness and phonotactic learning', *Linguistic Inquiry*, 44(1), pp. 45–75.

Hayes, B. and Wilson, C. (2008) 'A maximum entropy model of phonotactics and phonotactic learning', *Linguistic Inquiry*, 39(3), pp. 379–440.

Myers, J. and Tsay, J. (2015) 'Mandarin Wordlikeness Project [Raw data]'. Available at: http://lngproc.ccu.edu.tw/MWP/index.html.