Loan Phonology

EDITED BY

Andrea Calabrese
W. Leo Wetzels

LOAN PHONOLOGY

AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE

General Editor

E.F.K. KOERNER

Zentrum für Allgemeine Sprachwissenschaft, Typologie
und Universalienforschung, Berlin
efk.koernan@rz.hu-berlin.de

Series IV – CURRENT ISSUES IN LINGUISTIC THEORY

Volume 307

Andrea Calabrese and W. Leo Wetzels (eds.)

Loan Phonology

LOAN PHONOLOGY

Edited by

ANDREA CALABRESE
University of Connecticut, Storrs

W. LEO WETZELS
Université de Paris III-Sorbonne Nouvelle/
LPP, CNRS & VU University Amsterdam

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

Table of contents

Foreword

The articles contained in this volume originate from two sources. Manuscripts by Andrea Calabrese, Michael Friesner, Miren Lourdes Oñederra, and Lori Repetti are the result of the Going Romance XX loan phonology workshop. The other contributions are written by linguists specializing in this field who the volume editors invited.

We would like to take this opportunity to thank those colleagues who provided a helping hand in reviewing the papers, the CILT series editor, as well as the members of the CILT advisory editorial board for their useful comments. We would also like to extend our gratitude to those institutions which financially supported Going Romance XX: the Royal Netherlands Academy of Arts and Sciences (KNAW), the Faculty of Humanities at the VU University, the Algemeen Steunfonds of the VU University, the Leiden University Centre of Linguistics (LUCL), the Faculty of Humanities at the University of Amsterdam, the Utrecht Institute of Linguistics, and the Faculty of Humanities at Radboud University Nijmegen.

May 2009

Andrea Calabrese
Leo Wetzels

Loan phonology

Issues and controversies

[1]Andrea Calabrese & [2]W. Leo Wetzels
[1]University of Connecticut/[2]Université de Paris III–Sorbonne Nouvelle/LPP,
CNRS & Vrije Universiteit Amsterdam

The past decade has been characterized by a great interest among phonologists as to how the nativization of loanwords occurs. The general consensus is that loan-word nativization provides a direct window for studying how acoustic cues are categorized in terms of the distinctive features relevant to the L1 phonological system as well as for studying the true synchronic phonology of L1 by observing its phonological processes in action. The collection of essays in this volume provides an overview of the complex issues phonologists face when investigating this phe-nomenon and, more generally, the ways in which unfamiliar sounds and sound sequences are adapted to converge with the sound pattern of the native language.

Speakers borrow words from other languages to fill gaps in their own lexical inventory. The reasons for such lexical gaps vary greatly: cultural innovation may introduce objects or actions that do not have a name in the native language; native words may be perceived as non-prestigious; names of foreign cities, institutions, and political figures which were once unknown may have entered the public eye; new words may be introduced for play, etc.

Word borrowing can occur under two different scenarios. In the first, the borrowing may be implemented by a bilingual speaker that fills a gap in one of the languages he knows, L1, the recipient language, by taking a word from the other language he knows, L2, the donor language. In this case, the usual assump-tion (but see Footnote 1 below, for an alternative) is that the speaker retrieves the underlying representation of the borrowed word from his mental dictionary (the long-term memory storage for lexical items) for L2 and generates its surface repre-sentation while speaking L1. If the surface representation of the word is generated by using the phonological, or more generally, the grammatical system of L1, the word undergoes adaptations and adjustments and is nativized according to the grammar of L1.[1] We will call this event nativization-through-production.

---

1. The alternative is that the surface representation of the word is generated by using the L2 grammatical system. In this case, the word would be pronounced in its proper L2 shape.

In the other scenario, the borrowing is implemented by a speaker that fills a gap in his language by taking a word from another language he knows poorly or not at all.[2] In this case he needs to learn the relevant word. Once the learned word is uttered publicly or even silently by the speaker to himself, it is a loanword. Given that the speaker does not speak the second language well, the word will display adjustments and adaptations. The hypothesis is that these modifications have already occurred during perception and learning. One can call this scenario nativization-through-perception.[3]

These two scenarios essentially correspond to the two current models of loanword phonology: one essentially assumes that borrowing occurs only in the nativization-through-production scenario; Paradis & Tremblay (this volume) call it the phonological stance model (Hyman 1970; Danesi 1985; LaCharité & Paradis 2005; Paradis & LaCharité 1997; Paradis & Prunet 2000; Jacobs & Gussenhoven 2000; see also Paradis & Tremblay [this volume]).[4] The other model essentially assumes that borrowing occurs only in the nativization-through-perception scenario, referred to by Paradis & Tremblay (this volume) as the perceptual stance model (Silverman 1992; Yip 1993; Kenstowicz 2003b; Peperkamp & Dupoux 2002, 2003; see also the articles by Boersma & Hartman, Kim, and Calabrese in this volume).

The crucial difference between the two models has to do with the input to the nativization process. According to the perceptual stance model, it is the acoustic signal produced by the surface phonetic representation of the word; in contrast, the phonological stance model assumes that it is an abstract long-term memory

---

2.   Observe that this situation is the usual one for speakers of indigenous languages during the first stages of contact with the official language, as in the South-American native communities or the aboriginal communities of Papua New Guinea, etc.

3.   Another possibility, most recently discussed by Jacobs & Gussenhoven (2000), is that during perception and learning, the acoustic representations of the non-native segments are faithfully mapped into abstract featural representations, which are then encoded in long-term memory. These faithful featural representations of L2 sounds may obviously contain feature combinations that are characteristic of L2 and not allowed in L1. When this occurs, these feature combinations are modified during production in L1. It is, however, unlikely that such a faithful acquisition of non-native segments is ever possible. Current research starting from Dupoux et al. (1999; but see also Polivanov 1931) demonstrates that all types of modifications of non-native segments and words already occur in perception, which is heavily influenced by L1 grammatical categories.

4.   Calabrese (1988, 1995) and Connelly (1992) adopted a similar perspective in their analysis of loanword nativization.

(i.e., underlying) representation.[5] Another difference between the two models involves the nature of the nativization process: according to the phonological stance model, nativization is by force phonological insofar as the surface shape of the loanword is generated by the phonology of the recipient language. For the perceptual stance model, nativization can be both phonetic and phonological, as discussed below.

This book provides the reader with a collection of works representative of these two models. The phonological stance model is represented by the article "Nondistinctive Features in Loanword Adaptation: The unimportance of English aspiration in Mandarin Chinese phoneme categorization" by Carole Paradis and Antoine Tremblay. It investigates the treatment of stops in loanwords from English into Mandarin Chinese. As mentioned above, the phonological stance model proposes that nativization is brought about by the phonological processes characterizing speech production. According to this view, as earlier formulated by LaCharité & Paradis (2005), adapters always start with underlying representations of L2 words containing the L2 segments, because the adapters are bilingual in L1 and L2. The input to the adaptations is, therefore, always an abstract morphophonemic representation of the L2 word. Repairs to the L2 segments or strings are implemented so as to avoid the production of structures that are illicit in L1. Therefore, speakers should adapt loanwords by operating on a phonological/phonemic level that abstracts away from the details of allophonic and phonetic realization.

Mandarin Chinese (MC) distinguishes voiceless aspirated from voiceless unaspirated stops, yet dominantly adapts both phonetically aspirated (as in 'pie') and unaspirated voiceless stops ('spy') from English as aspirated in MC. Although all voiceless stops in English, regardless of whether they are aspirated or not, systematically yield an aspirated stop in MC, voiced English stops always result in unaspirated MC ones. Therefore, it appears that English stop aspiration, which is phonetic, does not influence phoneme categorization in MC, in spite of the fact that MC has phonemic aspirated stops. In other words, even though their native language predisposes MC speakers to distinguish aspirated from unaspirated stops, they appear not to rely on aspiration/nonaspiration in English to determine phoneme categorization in MC. According to Paradis & Tremblay, this provides evidence against the perceptual stance in loanword phonology which maintains

---

5.   Obviously this would be possible only for fully bilingual individuals. It follows that bilingual speakers play a fundamental role in the generation of loanwords. This is not to say that the phonological model denies that borrowers have access to the surface L2 representations. See, for example, LaCharité & Paradis (2005), who discuss adaptations based on 'naïve phonetic approximation', and who distinguish between 'naive phonetic approximation' and 'intentional phonetic approximation', where new phonemes are introduced into L1. See also the discussion (Conclusion) of Tremblay & Paradis' contribution to this volume.

that crucial information regarding loanword adaptation is phonetic; instead such data supports the phonological stance, according to which distinctive information exclusively is relevant to loanword adaptation.

As discussed previously, in the perceptual stance model the input to the adaptations is a surface phonetic representation of the L2 word and the nativization process occurs during perception when the new words are learned. The models that adopt this scenario can be divided into two groups. According to the first group, the adaptations observed in loanword nativization are accounted for by processes particular to perception and are fundamentally based on the notion of phonetic approximation/similarity. As for the other group, the adaptations involve the same phonological processes that characterize speech production.

The models assuming that nativization occurs in perception and are based on phonetic approximation/similarity can be traced back to Hermann Paul (1880). In his discussion of loanword phonology, he hypothesized that a host speaker, upon encountering a foreign segment, matches this phonetic signal with the native segment with which it is most closely related. Paul implicitly assumed that this match involves a perceptual similarity judgment based on *Sprachgefühl*, the feeling of language: speakers adapt a non-native segment to one which they 'feel' most closely resembles the former acoustically.

The models of loanword phonology that employ acoustic/perceptual similarity as the basis for the treatment of the loanwords (Silverman 1992; Yip 1993; Kenstowicz 2003a, b; Peperkamp & Dupoux 2003) develop this traditional view. According to them, the replacement operation between the non-native and the native segment is strictly based on phonetic similarity between the outputs of the donor and recipient languages. For example, according to Peperkamp & Dupoux (2003), the equivalences in loanword adaptation are based on a similarity that is defined as "acoustic proximity or proximity in the sense of fine-grained articulatory gestures."

"Mandarin Adaptations of Coda Nasals in English Loanwords" by Feng-Fan Hsieh, Michael Kenstowicz and Xiaomin Mou in this volume argues for such a perceptual model. This article is an investigation of the adaptation of English VN rhymes into Mandarin Chinese. The adaptation of the coda nasal is determined by the position of the vowel in the source word on the front-back, second formant (F2) dimension. Thus, the front vs. back quality of the vowel in English determines the substitution as [n] or [ŋ], respectively. When the vowel occupies a medial position on this dimension, as in the case of [ʌ] or schwa, the place of articulation of the English nasal coda is largely preserved. The consequence is that in the vowel + nasal consonant sequences, the vowel, which is phonetically more salient, determines the direction of adaptation, not the phonemically contrastive nasal itself, despite the fact that in MC the vowel differences heard in the source language are allophonic, not phonemic.

Together with the other articles of this collection, this work provides robust evidence demonstrating that the input to the adaptation in loanwords is phonetic. Most other articles in this volume reach the same conclusion and thus hypothesize that loanword nativization occurs during perception, although they also argue that the adaptations evident in loanwords are phonological in nature. This is the case, for example, of "Loanword Adaptation as First-Language Phonological Perception" by Paul Boersma and Silke Hamann, who argue that loanword adaptations can only be truly understood in terms of the perception seen as an active process involving the mapping from raw sensory data to a more abstract mental representation. According to these authors, this process is fully phonological and involves an Optimality Theory (OT) interaction between structural and cue constraints. The structural constraints that play a role in a given language perception are the same ones active in production. In both perception and production, these constraints are ranked high. In perception, however, they interact not with faithfulness constraints, as they do in production, but with cue constraints. Cue constraints evaluate the relation between the input of the perception process (the auditory-phonetic form) and the output of the perception process (the phonological surface form). The result is that the satisfaction of these structural constraints in perception typically leads to processes different from those that occur in production.

Articles by Hsieh, Kenstowicz & Mou, and Boersma & Hamann are couched within the OT model of phonology, as are many other works adopting the perceptual stance. Adoption of OT is, however, not required to pursue the idea that loanword phonological adaptations occur in perception. The article "Korean Adaptation of English Affricates and Fricatives in a Feature-Driven Model of Loanword Adaptation" by Hyunsoon Kim, who does not reference OT, in fact also proposes that the perceived acoustic properties of L2 are structured according to the phonological categories of L1, specifically, according to L1 distinctive features and syllable structure, rather than in terms of the unstructured L2 acoustical input per se or of L2 phonological categories. In this model, it is assumed that acoustic parameters and cues are extracted in the first stage of L1 perception and that they are mapped into L1 linguistic entities such as distinctive features and syllable structure in conformity with the L1 grammar. In this way, loanwords are extracted and stored in a mental lexicon where each word is represented as a sequence of syllabified distinctive feature bundles stored in long-term memory.

Another article that also investigates loanwords in the context of speech perception but does not adopt OT is Andrea Calabrese's "Perception, Production and Acoustic Inputs in Loanword Phonology". He also investigates how a learner constructs mental representations of L2 sounds and structures by means of complex inferential computations. In this process, the learner adjusts these non-native sounds and structure so as to make them familiar, and therefore 'understand'

them accordingly in perceptual mental representations. An important concern for Calabrese is that, if perception of new words involves interpretation and inferential computation, it loses its primary function of tracking external reality, the environment; it becomes detached from reality and prone to illusions. He proposes that listeners always have direct access to the acoustic signal through a representation that is stored in a short-term acoustic working memory buffer, 'echoic memory' (see Neisser 1967). Although illusion-like, interpretative failures may occur, the acoustic representations preserved in echoic memory tie perception to external reality.

The issue of the construction of the underlying representation (UR) of loanwords is also the main focus of the articles by Nevins & Braun and Wetzels. These URs can be very abstract and quite different from the L2 URs. In his article "Nasal Harmony and the Representation of Nasality in Maxacalí: Evidence from Portuguese Loans", Leo Wetzels argues that nasality in the Brazilian indigenous language Maxacalí is contrastive only in the case of vowels. Nasal consonants are always derived by spreading the nasal feature of the vowel onto its syllable onset and coda if there is one. Wetzels shows that in Brazilian Portuguese (henceforth: BP) loanwords to Maxacalí, the original nasal onsets of the BP words are analyzed as being the outcome of this spreading rule. As he puts it, "In other words, confronted with a BP syllable containing a nasal onset and an oral nucleus, the speaker of Maxacalí interprets the nasal onset as an indication of the nasality of its nucleus." Therefore, faced with BP words such as *carneiro* 'sheep' [kahˈneɾu], a Maxacalí speaker postulates a UR where the nasality is a property of the vowel /kahDẽT/. The rule then spreads the nasality onto the preceding onset voiced stop and the following coda [kahnẽn]. If the vowel is interpreted as oral in the borrowing, its onset is also nonnasal, as expected if nasality is a property of the vowel and the partial nasality in word-initial oral syllable onsets is derived by rule, cf. Maxacalí [ᵐbahtet] from BP [mahˈtelu] *martelo* 'hammer'

Awareness of the rules and constraints of the L1 grammar, therefore, leads to the postulation of more abstract representations for L2, in particular the postulation of a representation for some L2 word consistent with the rules and constraints of L1. The paper "The Role of Underlying Representations in L2 Brazilian English" by Andrew Nevins and David Braun discusses the pronunciation of English as pronounced by Brazilians (Brazilian Portuguese English, BPE). Brazilian Portuguese has a rule changing the rhotic /r/ to a laryngeal fricative in word initial position: [dȝiˈɾɛtu] *direto* 'straight' vs. [ˈhɛtu] *reto* 'straight on'. Interestingly, in their pronunciation of English, Brazilian speakers replace word-initial /h/ with [r] (e.g., [rom] (or [hom]) for *home*). Nevins & Braun explain this replacement by hypothesizing that when exposed to English words, a Brazilian learner observes that the rule debuccalizing [r] into [h] does not apply to English. When faced to word-initial

/h/ in English, he hypothesizes that it derives from underlying /r/ as in his own language. Given that he has postulated that r-debuccalization does not apply in English, this hypothesized /r/ surfaces in the English word as can be seen in [rom] for *home*. Therefore the speaker postulates a UR consistent with the phonological system of L1.

The conclusion in most of the papers in this collection is that the nativization of loanwords occurs under the nativization-through-perception scenario, i.e., when the L2 words are perceived and learned. This is again shown with another aspect of BPE pronunciation discussed by Nevins & Braun: the affrication of coronal stops before the vowel [u]. The authors relate this unexpected process to the fact that /u/ is fronted after coronals English. In their analysis, this fronted /ü/ becomes the diphthong [iu]; the high front vocalic component manages to trigger the affrication characteristic of their native BP phonology.

Nevins & Braun show that in order to account for the borrowing of allophonic [ü], the input must be phonetic and not phonemic as assumed by the phonological stance model. Simultaneously, the assumption that phonetic similarity is essential to the adaptation found in loanword phonology as hypothesized by some perceptual stance theorists cannot account for the affrication we find in this case in BPE. Crucially, the adaptation of the loanword must be phonological in nature.

The converging evidence is that, if one assumes that the adaptations are indeed phonological, one could reinterpret the cases for the phonological stance model in terms of the perceptual stance model as involving an alternative phonological treatment of the acoustic input, without requiring bilingualism and access to abstract underlying forms of L2. Clearly, if the Mandarin Chinese adaptors possess full mastery of both the phonological grammars of English and MC, they 'know' that in English aspirated and non-aspirated stops are in complementary distribution, i.e., belong to the same phonological class. Their choice of the feature [aspirated] as the relevant corresponding lexical feature in MC may be directly imposed by the English grammar, if Iverson & Salmons' (1995) hypothesis that [aspirated] is the underlying feature for voiceless stops is correct. Otherwise, perhaps their choice for [aspirated] as the generalized feature owes to the observation that aspiration is realized in the perceptually more salient stressed syllables in the English loans. If, on the other hand, no knowledge of the English phonology could be assumed, one would need to explain why the English surface system [p, pʰ, b] is categorized in terms of the MC distinctive categories the way it is. In other words, although Paradis & Tremblay (this volume) convincingly show that the perceptual stance model alone is inadequate for predicting the MC nativization of the English laryngeal features, it is also clear that, in the case of a monolingual MC speaker, perception would have a role in explaining why [p, pʰ] (> MC /pʰ/) are classified together as a single phonological class and

as separate from [b] (> MC /p/). One possibility is that the feature [voice] is to be rejected, as proposed by Halle & Stevens (1971) and that the distinction between aspirated and non-aspirated (lenis) consonants in MC is to be made in terms of the features [stiff vocal cords] and [slack vocal cords], as the latter authors propose for Korean. This would yield a classification of MC aspirated "voiceless" obstruents as [+ stiff vocal cords, – slack vocal cords, + spread glottis], unaspirated "voiceless" obstruents [– stiff vocal cords, – slack vocal cords, – spread glottis]. In the case of the nativization of English stops in Mandarin Chinese one could then propose the following: given that [+stiff vocal cords] stops are always aspirated in MC, we need a rule such as [+stiff vocal cords] → [+spread glottis]/[___, –sonorant]. One can then propose that during perception all voiceless stops are interpreted by the MC learner according to that rule, so that the allophonic distribution of [spread glottis] in English is overridden. It is unclear if there is phonetic evidence for this hypothesis, but it is obvious that for the monolingual MC adaptor, there would have to be some acoustic property shared by English [p^h, p], which is lacking in English [b], in order to explain the classification he is making. The difference between bilingual and monolingual adaptors therefore becomes very relevant.

At this point it may be tempting to simply assume that all nativization occur during perception, though this would be an implausible conclusion. Anyone familiar with bilingual environments knows that nativized loanwords can be innovatively produced by bilinguals simply by taking one word from one of the languages they know and adapting it into the other language they know, e.g., an English-Italian bilingual may take the English word for *street* and adapt it into Italian [stritta]. Still, there remains the issue of the overwhelming evidence supporting the observation that for the majority of loanwords, the input seems to be the L2 word in its surface phonetic representation.

A possible solution to this problem is suggested in "The Adaptation of Romanian Loanwords from Turkish and French" by Michael L. Friesner, who examines several factors affecting loanword adaptation, using a data set of Romanian loanwords from Turkish and French. It proposes that in order to get a full picture of how loanwords are nativized, one must consider not only different modules, such as the phonology and the morphology, but also different levels, including linguistic differences and external explanations such as orthography and, most importantly, social factors. For example, there is a striking difference in the nativization of loanwords from Turkish and French into Romanian with regard to gender. Whereas the gender was assigned to Turkish words arbitrarily, this was not so with French, where the gender of the word proved pertinent. This is because French borrowings were usually facilitated by scholars who had learned French grammar formally and thus had a greater awareness of the gender of French words. There was also a

need to have these words fit into a native pattern, since French words were often borrowed out of a conscious effort to 're-Latinize' the language.

Thus, socio-political factors have an impact on the nativization patterns. Suppose that for normative social reasons, the input to nativization even during production is always a surface word as it is 'heard' and not its abstract mental representation. This is because words are accepted in the linguistic community in their surface shape, which thus acquires a normative status. The abstract UR is used only to pronounce the L2 word correctly although it can be pronounced with an accent, and not as input to nativization. In this nativization scheme, a bilingual borrower first produces the word in L2 and then uses that surface representation as input to the nativization process, which is phonological. If this is correct, the perceptual stance and phonological stance models no longer need be contrasted, and could be largely unified: the input to nativization is always phonetic, the word as it is "heard". The treatment, on the other hand, is always phonological and it can occur either during perception or during production.

The importance of the social dimension of language in determining the properties of loanwords is also discussed in the article "Early Bilingualism as a Source of Morphonological Rules for the Adaptation of Loanwords: Spanish loanwords in Basque" by Miren L. Oñederra. She considers the special situation circumscribing Spanish loanwords into Basque: Basque speakers are bilingual (some simultaneous, some sequential) with respect to Spanish, and have been for many years, with the result that once phonologically natural processes of substitution have become defunctionalized and institutionalized into synchronically arbitrary patterns. This study demonstrates the intertwining influences between linguistically unrelated yet socially coexistent languages over a long period of time, underscoring how contact this close can result in the loss of phonological motivation for some of the sound substitutions that occur as one language incorporates words from the other.

Finally, the complexities of the nativization process are the subject of Lori Repetti's article "Gemination in English Loans in American Varieties of Italian". It deals with the process whereby a singleton consonant in the loaning language is adapted as a geminate consonant in the borrowing language. This process is very common cross-linguistically and is attested in Japanese, Finnish, Kannada, Maltese Arabic, Hungarian, and Italian (including North American varieties), as well as many other languages. Repetti argues that a combination of factors is needed to account for gemination in loanwords, e.g., lexical considerations, morpho-phonological constraints, and, importantly, perceptual factors. This again demonstrates that perception and production cannot be separated in the study of nativization in loanwords, but must be always seen in their synergetic interaction. This is what we believe to be the most important conclusion of this collection.

References

Calabrese, Andrea. 1988. *Towards a Theory of Phonological Alphabets*. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, Mass.

Calabrese, Andrea. 1995. "A Constraint-Based Theory of Phonological Markedness and Simplification Procedures". *Linguistic Inquiry* 26:2. 373–463.

Connelly, Mark. 1994. *Phonological Markedness and Second Language Error Interpretation,* Ph.D. dissertation. Harvard University, Cambridge, Mass.

Danesi, Marcel. 1985. *Loanwords and Phonological Methodology*. Ville LaSalle, Quebec: Didier.

Dupoux, Emmanuel, Kazuhiko Kaheki, Yuki Hirose, Christophe Pallier, Stanka Fitneva, & Jacques Mehler. 1999. "Epenthetic Vowels in Japanese: A perceptual illusion". *Journal of Experimental Psychology: Human perception and performance* 25. 1568–1578.

Hyman, Larry. 1970. "The Role of Borrowings in the Justification of Phonological Grammars". *Studies in African Linguistics* 1. 1–48.

Iverson, Gregory K. & Joseph C. Salmons. 1995. "Aspiraton and Laryngeal Representation in Germanic". *Phonology* 12. 369–396.

Jacobs, Haike & Carlos Gussenhoven. 2000. "Loan Phonology: Perception, salience, the lexicon and OT". *Optimality Theory: Phonology, syntax, and acquisition* ed. by Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer, 193–210. Oxford: Oxford University Press.

Kenstowicz, Michael. 2003a. "Salience and Similarity in Loanword Adaptation: A case study from Fijian". *Language Sciences* 29. 316–340.

Kenstowicz, Michael. 2003b. "The role of Perception in Loanword Phonology: A review of *Les emprunts linguistiques d'origine européenne en Fon by Flavien Gbéto*". *Studies in African Linguistics.* 32. 95–112.

LaCharité, Darlene & Carole Paradis. 2005. "Category Preservation Versus Phonetic Approximation in Loanword Adaptation". *Linguistic Inquiry* 36:2. 223–258.

Neisser, Ulric. 1967. *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Paradis, Carole & Darlene LaCharité. 1997. "Preservation and Minimality in Loanword Adaptation". *Journal of Linguistics* 33. 379–430.

Paradis, Carole & Jean-François Prunet. 2000. "Nasal Vowels as Two Segments: Evidence from borrowings". *Language* 76. 324–357.

Paul, Hermann. 1880. *Prinzipien der Sprachgeschichte*. Halle/Saale: Max Niemeyer.

Peperkamp, Sharon & Emmanuel Dupoux. 2002. "A Typological Study of Stress 'Deafness'". *Laboratory Phonology* 7 ed. by Carlos Gussenhoven & Natasha Warner, 203–240. Berlin: Mouton de Gruyter.

Peperkamp, Sharon & Emmanuel Dupoux. 2003. "Reinterpreting Loanword Adaptations: The role of perception". *Proceedings of the 15th International Conference of Phonetic Sciences.* Barcelona: Causal Productions. 367–370.

Polivanov, Evgenij D. 1931. "La perception des sons d'une langue étrangère". *Travaux du Cercle Linguistique de Prague* 4. 79–96. [English translation, "The Subjective Nature of the Perceptions of Language Sounds", *Selected Works: Articles on general linguistics*, 223–237. The Hague: Mouton. 1974.]

Silverman, Daniel. 1992. "Multiple Scansions in Loanword Phonology: Evidence from Cantonese". *Phonology* 9. 289–328.

Yip, Moira. 1993. "Cantonese Loanword Phonology and Optimality Theory". *Journal of East Asian Linguistics* 2. 261–291.

Loanword adaptation as first-language phonological perception*

Paul Boersma & Silke Hamann

We show that loanword adaptation can be understood entirely in terms of phonological and phonetic comprehension and production mechanisms in the first language. We provide explicit accounts of several loanword adaptation phenomena (in Korean) in terms of an Optimality-Theoretic grammar model with the same three levels of representation that are needed to describe L1 phonology: the underlying form, the phonological surface form, and the auditory-phonetic form. The model is bidirectional, i.e., the same constraints and rankings are used by the listener and by the speaker. These constraints and rankings are the same for L1 processing and loanword adaptation.

Figure 1 shows a simplified version of an existing model for first-language (L1) processing (Boersma 1998, 2000, 2007ab).[1] The model is bidirectional, i.e., it accounts for the behaviour of the listener (on the left) as well as the speaker (on the right). In both directions, processing is assumed to be handled by the interaction of Optimality-Theoretic constraints.

Phonological production (top right) is described in terms of an interaction between structural and faithfulness constraints (McCarthy & Prince 1995). Perception (bottom left) is described in terms of an interaction between structural and cue constraints (Boersma 2000, 2007ab). The remaining two processes, word recognition (top left) and phonetic implementation (bottom right), are (in this simplified version) described by one set of constraints each (faithfulness and cue constraints, respectively).

---

1.   We explain some simplifications in footnotes. One simplification is that a more elaborate model (Boersma 1998, 2007ab; Apoussidou 2007) requires additional representations, such as an articulatory form (below the auditory-phonetic form in Fig. 1) and a morpheme level (above the underlying form).

COMPREHENSION PRODUCTION
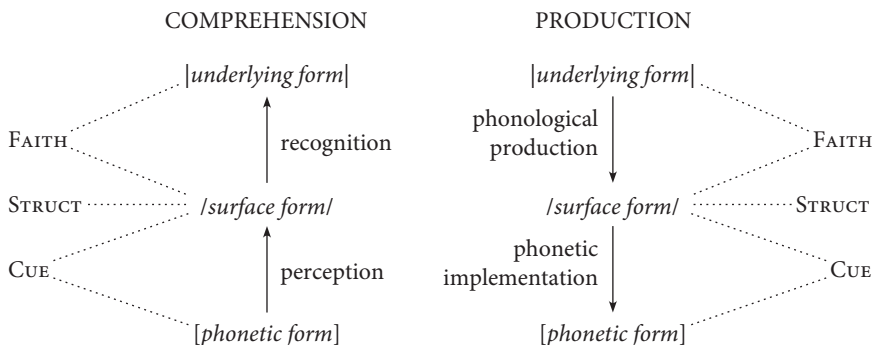
Figure 1. A single model for L1 processing as well as loanword adaptation

The roles of all the ingredients of the model in Fig. 1 will become clear in our discussion of the examples that follow. The idea to take home from Fig. 1 is that structural constraints play a role both in production and in comprehension, although they interact with different constraints in these two directions of L1 processing. We will show that the L1 model of Fig. 1 suffices to account for many loanword adaptation phenomena, thereby doing away with the loanword-specific devices that have appeared in other (earlier as well as later) proposals in the literature.

1. Superficial differences between Korean native phonology and loanword adaptation

Our first subject of discussion is the often commented fact that a process superficially describable as *vowel insertion* is much more common in loanword adaptation than in native phonologies. As our example in this paper, we analyze observations about vowel insertion in English loanwords in Korean (H. Kang 1996, 1999; Y. Kang 2003; Kabak 2003).

Illicit surface structures seem to be handled differently in the native Korean phonology than in English-to-Korean loanword adaptation. In native Korean phonology, such structures are typically avoided by processes of neutralization, assimilation, and deletion, but never by vowel insertion. The underlying form |patʰ| 'field' is produced as the surface structure /.pat./, an underlying |os| 'clothes' is produced as surface /.ot./, an underlying |kaps| 'price' as /.kap./, and an underlying |kuk+min| 'nation' as /.kuŋ.min./. According to all authors, the avoidance of the faithful */.patʰ./ is due to a Korean structural restriction against aspirated codas, the avoidance of the faithful */.os./ is due to a Korean structural restriction

against strident codas, the avoidance of the faithful */.kaps./ (or */.kapt./) is due to a Korean structural restriction against coda clusters, and the avoidance of the faithful */.kuk.min./ is due to a Korean structural (phonotactic) restriction against segmental sequences like */km/. Crucially, all eight constraints involved here (faithfulness for aspiration, faithfulness for stridency, segmental faithfulness, faithfulness for manner, and the four structural constraints) could have been satisfied by inserting a vowel (/.pa.tʰɨ./, /.o.sɨ./, /.kap.s'ɨ./, /.ku.kɨ.min./), but this is not what Korean speakers do. Apparently, the faithfulness constraint against surface vowels that have no correspondent in the underlying form (i.e. the constraint DEP-V), is ranked quite high in native Korean phonology.

At the same time, however, the adapted English words *deck*, *mass*, *false* and *picnic* can show up as /.tɛ.kʰɨ./, /.mæ.s'ɨ./, /.pʰol.s'ɨ./ and /.pʰi.kʰɨ.nik./, respectively, i.e. with apparently inserted vowels. For a 'minimal view' of loanword adaptation, this poses a problem. Under such a minimal view, learners would first store the English surface forms as the segmentally closest Korean underlying forms |tɛkʰ|, |mæs'|, |pʰols'| and |pʰiknik|, and then run these underlying forms through the native Korean constraint ranking. If this were correct, the four words would have to show up as /.tɛk./, /.mæt./, /.pʰol./ and /.pʰiŋ.nik./, but this is not what happens.[2] All OT analyses therefore agree (as do we) that this minimal close-copy-plus-L1-filtering is not how loanword adaptation proceeds. Apparently, loanword adaptation is either performed in production with a different constraint ranking than the native phonology (e.g. with a low-ranked DEP-V), or the underlying forms of loanwords are not stored as close copies of the surface forms of the donor language (because the native Korean perception process changes the form first).

Both of these possibilities have been considered in the literature. All the production-based accounts have to invoke loanword-specific mechanisms, such as loanword-specific rankings or loanword-specific constraints. However, all the perception-based accounts that do not assume the three-level model of Fig. 1 have to invoke direction-specific rankings or constraints. In §2 to §5 we analyze the Korean facts in the three-level L1-only framework of Fig. 1, showing that our analysis does not have to invoke any loanword-specific mechanisms and works solely with rankings and constraints that are the same for speakers and listeners. In §6 we discuss previous analyses found in the literature and show why these fail to work when not assuming loanword- or direction-specific mechanisms. In §7 we discuss some interesting additional issues.

---

2.   Forms like these, i.e. without vowel epenthesis, sometimes do occur; we discuss them in §4.3 and Footnote 11.

2.   Native Korean phonological processes: No vowel insertion

Here we show in detail how the three processes of Korean phonological produc-
tion mentioned in §1 work, and give an Optimality-Theoretic account that will
lead us to establish a ranking in which Dep-V must be ranked high.

2.1   An L1 phonological process: Neutralization

One way to satisfy Korean structural restrictions is to neutralize a featural con-
trast. Korean plosives, for instance, come in three manners: lax (/t/), aspirated
(/tʰ/), and fortis (/t'/). We denote them by the feature combinations /−tense,−asp/,
/+tense,+asp/, and /+tense,−asp/, respectively (Iverson 1983; H. Kim 2005). In codas,
all plosives surface as lax, i.e. any underlying |+tense| is turned into /−tense/
and any underlying |+asp| is turned into /−asp/. The Korean word meaning
'field', for instance, is underlyingly |patʰ|, as evidenced by the locative /.pa.tʰɛ./
'in the field', from underlying |patʰ+ɛ|. In final position, the underlying |patʰ| is
produced as the surface form /.pat./, with a laryngeal neutralization that can be
described in terms of an interaction between structural and faithfulness con-
straints. We write the structural constraint against aspirated codas as */+asp ./.
This structural constraint must outrank a faithfulness constraint for underly-
ing aspiration, e.g. Ident(asp). Tableau (1) gives the interaction (after H.Kang
1996; also Y.Kang 2003:224).

(1)   L1 Korean production: coda deaspiration

|patʰ|	*/+asp ./	Dep-V	Max-C	Ident(asp)	*/C ./
/.patʰ./	*!				*
☞ /.pat./				*	*
/.pa.tʰɨ./		*!			
/.pa./			*!		

Crucially, we see that Dep-V has to be ranked quite high: in order that aspiration
faithfulness cannot force insertion of an epenthetic vowel, Dep-V has to outrank
Ident(asp); and in order that, say, a general constraint against codas cannot force
vowel epenthesis, Dep-V has to outrank the structural constraint */C ./. We also
see that the faithfulness constraint Max-C ("an underlying consonant should have
a correspondent in the surface form") has to outrank both Ident(asp) and */C ./.
The ranking of Dep-V above Max-C is explained in §2.2.

Coda neutralization is not restricted to laryngeal features. The Korean word
meaning 'clothes', for instance, is underlyingly |os|, as evidenced by the nomi-
native /.o.si./, from underlying |os+i|. In non-prevocalic position, |os| surfaces

as /.ot./. This strident neutralization can be described in terms of an interaction between a structural constraint against strident segments in coda position, */+stri ./, and a faithfulness constraint for underlying strident specifications, IDENT(stri), as in (2).

(2)   L1 Korean production: strident neutralization

|os|	*/+stri ./	Dep-V	Max-C	Ident(stri)	*/C ./
/.os./	*!				*
☞ /.ot./				*	*
/.o.sɨ./		*!			
/.o./			*!		

As in Tableau (1), we see a high ranking of Dep-V: this constraint has to outrank Ident(stri) so that the latter cannot force insertion of an epenthetic vowel.

2.2   Another L1 phonological process: Deletion

Another way to satisfy Korean structural restrictions is to delete a consonant. Korean codas can have two consonants underlyingly, but only one will surface. For instance, the Korean word meaning 'price' is underlyingly |kaps|,[3] as evidenced by the form /.kap.sʼi.sʼa.ta./ 'cheap', from underlying |kaps+i sʼata|. In final position, the underlying form |kaps| is produced as the surface form /.kap./, with a deletion that can be accounted for in terms of an interaction between Max-C and the structural constraint */CC ./ ("no complex codas"). This is shown in Tableau (3).

(3)   L1 Korean production: final consonant deletion

|kaps|	*/+stri ./	*/CC ./	Dep-V	Max-C	Ident(stri)	*/C ./
/.kaps./	*(!)	*(!)				*
/.kapt./		*!			*	*
/.kap.sɨ./			*!			*
☞ /.kap./				*		*
/.ka./				**!		

Again, we see that DEP-V is ranked high. Here (unlike in §2.1), the ranking of DEP-V above MAX-C is crucial: it is better to delete a consonant than to insert a vowel.

2.3   A third L1 phonological process: Assimilation

The Korean noun meaning 'country' is underlyingly |kuk|, as evidenced by the form /.ku.kʌ./ 'Korean language'. Before nasal consonants, the form changes: an underlying |kuk+min| 'nation' is produced as the surface form /.kuŋ.min./. According to various authors (Iverson & Sohn 1994; Davis & Shin 1999), this change is due to the syllable contact law (Hooper 1976; Murray & Vennemann 1983; Vennemann 1988), which for Korean asserts that a coda should not be less sonorous than the following onset. Davis & Shin (also H. Kang 2002) therefore give an OT analysis in terms of an interaction of the structural constraint SYLLCON with various faithfulness constraints. As Davis & Shin notice, SYLLCON, DEP-V and MAX-C have to outrank faithfulness constraints for underlying sonority and/or nasality; a simplified version of their analysis is shown in Tableau (4).[4]

(4)   L1 Korean production: nasal assimilation

|kuk+min|	SYLLCON	DEP-V	MAX-C	IDENT(nas)	*/C ./
/.kuk.min./	*!				**
☞ /.kuŋ.min./				*	**
/.ku.kɨ.min./		*!			*
/.ku.min./			*!		*

Again we see a high ranking of DEP-V: in order that nasality faithfulness cannot force insertion of an epenthetic vowel, DEP-V has to outrank IDENT(nas).

2.4   A constraint ranking for native Korean phonological production

Together, the evidence from Tableaus (1) to (4) shows that DEP-V is high-ranked in native Korean production: it outranks at least four faithfulness constraints and one structural constraint.

---

4.   A candidate /.kuk.pin./, which violates the same constraints as the winner in (4), can be ruled out either by splitting IDENT(nas) into IDENT(son) and MAX(nas) (Davis & Shin 1999), or by realizing that IDENT(nas) could be ranked higher for underlying |+nas| segments than for underlying |−nas| segments, as an emergent result of frequency differences between |+nas| and |−nas| segments (Boersma 2008; cf. §5).
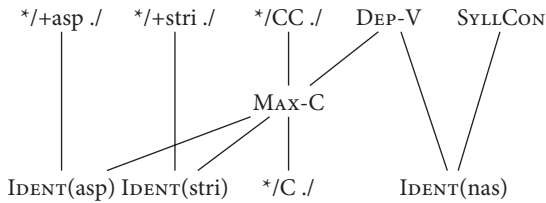
Figure 2. Crucial rankings for native Korean phonological production

The ranking in Fig. 2 again makes the point that vowel insertion is an avoided process in native Korean phonological production. In native Korean *perception*, the situation is rather different, as we show in the next section.

3. Native Korean perception of English sounds: Ubiquitous vowel insertion

In this section we make plausible that in their native perception processes, Korean listeners routinely insert vowels, and that this causes the perceptual insertion of vowels into auditory-phonetic forms of English. In this we follow Y. Kang (2003), who convincingly argues that Korean listeners of English insert vowels. Unlike Kang, however, we provide an Optimality-Theoretic formalization of this perception process. Following Boersma (1998), this formalization is done in terms of the three levels depicted in Fig. 1, i.e., the term 'perception' refers only to the mapping from an auditory-phonetic form to a phonological surface structure. Following Boersma (2000, 2007ab), we formalize perception in terms of an interaction between cue constraints and structural constraints: cue constraints evaluate the relation between the input of the perception process (the auditory-phonetic form) and the output of the perception process (the phonological surface form), while structural constraints evaluate only the output of this process.

We will see that the structural constraints that play a role in native Korean perception are the same ones that play a role in native Korean production (Fig. 2). In perception, they will again turn out to be ranked high, as in production (Fig. 2). In perception, however, they interact not with faithfulness constraints (as they do in production) but with cue constraints, and the result is that the satisfaction of these structural constraints will in perception typically lead to vowel insertion rather than to any of the three processes that occur in production (§2).

3.1 Korean perception of English segments: Cue constraints

We start our discussion of loanword adaptation with a discussion of foreign-language perception, because loanword adaptation must ultimately start from the

auditory-phonetic form (the sound) of the word in the donor language. In this section we illustrate how the L1-only model of Fig. 1 handles the Korean perception of English vowels and plosives. Our main point here is to show how in words like *tag* and *deck* Korean listeners insert a vowel, i.e. how they interpret them as /.tʰæ.kɨ./ and /.tɛ.kʰɨ./.

In a narrow phonetic transcription, the sounds of the English words *tag* and *deck* look like [_ᵗʰá̰ːɡ˥ ˳ɡ] and [_ᵈɛ̰k˥_ᵏʰ]. In these narrow auditory transcriptions, the underscore ("_") stands for the silence that occurs in plosives; "ᵗ" and "ᵏ" stand for the fortis alveolar and velar plosive release bursts, respectively; "ɡ" and "ᵈ" stand for the lenis velar and alveolar plosive release bursts; "ʰ" stands for the (English-type) moderately strong aspiration noise; "a̰" and "ɛ̰" are the IPA transcriptions for the two English front vowels; "ː" reflects the typical English lengthening of vowels before voiced consonants (Heffner 1937; House & Fairbanks 1953); "ɡ˥" and "k˥" stand for the formant transitions from a vowel into the velar stops; "ˊ" and "ˉ" stand for the high and mid-high fundamental frequency (F0) associated with English voiceless and voiced plosives in stressed syllables (House & Fairbanks 1953; Lehiste & Peterson 1961; Ohde 1984); and "˳" stands for the voicing murmur during the closure of a voiced plosive. All these concrete details are what English listeners use all day to make sense of their surrounding speech: they are the cues that English listeners use for interpreting the surrounding speech in terms of English-specific abstract phonological elements (features, segments, syllables). Together, these cues will lead an English listener to interpret the sounds [_ᵗʰá̰ːɡ˥ ˳ɡ] and [_ᵈɛ̰k˥_ᵏʰ] as the phonological surface structures /.tæɡ./ and /.dɛk./, where "." stands for a syllable boundary and e.g. the notation /t/ is a convenient shortcut for a more elaborate feature combination like [cor,−cont,−voi]. Importantly, auditory forms like [_ᵗʰá̰ːɡ˥ ˳ɡ] and [_ᵈɛ̰k˥_ᵏʰ] and surface forms like /.tæɡ./ and /.dɛk./ are representations that use different alphabets; the fact that our auditory and surface notations partially utilize some of the same symbols is purely coincidental.

When confronted with the sounds [_ᵗʰá̰ːɡ˥ ˳ɡ] and [_ᵈɛ̰k˥_ᵏʰ], a Korean listener will interpret the phonetic details in a different way from an English listener: a Korean listener will interpret these sounds in terms of Korean phonology. In this section we consider only the featural and segmental interpretations, leaving the interpretations in terms of syllable structure to §3.2, and phonotactically restricted interpretations to §3.3.

We start with the prevocalic English sounds [_ᵗʰ] and [_ᵈ]. We assume that a Korean listener will perceive both of them as a Korean alveolar plosive, i.e. as /t/, /t'/ or /tʰ/. In phrase-initial position, the plosives have the following pronunciations (Lisker & Abramson 1964; Han & Weitzman 1970; Hardcastle 1973;

Hirose, Lee & Ushijima 1974; Kagaya 1974; Cho, Jun & Ladefoged 2002): /to/ is pronounced as [_^{d}ò], with a lenis voiceless burst (i.e. a positive voice onset time, with possible slight aspiration) and a lowered F0 on the vowel; /t'o/ is pronounced as [_^{t}ó], with a fortis release burst (no aspiration) and a raised F0 on the vowel; and /t^{h}/ is pronounced as [_thó], with a fortis release burst, more aspiration noise than the English prevocalic /t/ has, and again with a raised F0. These differences in produced cues are reflected in the Korean perception of these three segments. When listening to initial plosives that vary in the degree of aspiration noise and in the height of F0, Korean listeners turn out to rely mainly on F0 to distinguish /t/ on the one hand (lowered F0) from /t'/ and /t^{h}/ on the other hand (raised F0); the distinction between /t'/ and /t^{h}/ is then made on the basis of aspiration noise (M.-R. Cho Kim 1994; Kim, Beddor & Horrocks 2002). Given these native Korean cue reliances, we can expect that Koreans interpret the plosives in the English sounds [_thó] and [_^{d}ō] as their phonemes /t^{h}/ and /t/, respectively. That they do this, has been confirmed in perception experiments (M.-R. Cho Kim 1994; Schmidt 1996; H. Park 2007) and is compatible with the loanword facts, as we will see.

Just asserting that English [_thó] and [_^{d}ō] tend to be classified by Korean listeners as /t^{h}o/ and /to/ does not suffice for our purposes: we need a formalization as well. Boersma (1997, 1998, 2000, 2006, 2007ab, 2008), Escudero & Boersma (2003, 2004), Escudero (2005), Boersma & Escudero (2008), Boersma & Hamann (2008), and Hamann (2009) provide such a formalization in terms of *cue constraints*. Just as faithfulness constraints do, cue constraints link two representations: whereas faithfulness constraints link underlying forms to phonological surface forms, cue constraints link auditory-phonetic forms to phonological surface forms (Fig. 1). The nature of cue constraints, though, is very different from that of faithfulness constraints: whereas faithfulness constraints link two discrete representations, cue constraints link a discrete representation (the surface form) to a continuously-valued representation, namely the auditory-phonetic form.

In order to establish the set of cue constraints for initial plosives, we have to establish first what the Korean representations look like. At the auditory-phonetic level, Koreans (just like the English) have universal representations like [_thó] and [_^{d}ō]. At the phonological surface level, we have used above the unitary phonemic symbols /t/, /t'/, and /t^{h}/, but these are just shorthands for the Korean-specific feature bundles /cor,−son,−cont,−tense,−asp/, /cor,−son,−cont,+tense,−asp/, and /cor,−son,−cont, +tense,+asp/, respectively. A relevant cue constraint, then, is *[^{h}]/−asp/, i.e. "moderately strong aspiration noise in the auditory form should not be perceived as the phonological feature value /−asp/ in the surface form". This constraint alone is enough to make sure that [_thó] is perceived as /t^{h}o/, as shown in Tableau (5).

(5)   Korean perception of the English initial /t/, i.e. the sound [_ᵗʰó] (first version)

[_ᵗʰó]	*[ʰ] /−asp/
/to/	*!
☞ /tʰo/	
/t'o/	*!

The perception tableau in (5) works as follows. The top left cell contains the input to perception, that is, the auditory-phonetic form [_ᵗʰó]. The three candidate cells contain the three candidate outputs of perception, i.e. the three phonological surface forms /to/, /tʰo/, and /t'o/. The first candidate violates the constraint *[ʰ]/−asp/, because the input phonetic form contains the sound [ʰ] and the output phonological surface form /to/ contains the feature value /−asp/. For the same reason, the third candidate also violates this constraint. As a result of the two violations, the listener cannot perceive [_ᵗʰó] as /to/ or /t'o/, and is left with no other option than to perceive [_ᵗʰó] as /tʰo/. The perception tableau in (5), then, has provided a formalization of what we earlier expressed in plain English.

We now turn to the perception of [_ᵈō], for which we have the same three candidate perceptions as for [_ᵗʰó]. The sound [_ᵈō] does not contain aspiration noise, so our old constraint *[ʰ]/−asp/ will not be able to distinguish between any of the three candidates. Instead, we can now use the counterpart of this constraint, which is *[*no noise*]/+asp/, i.e. "auditory absence of noisiness should not be perceived as the feature value /+asp/". As shown in Tableau (6), this constraint helps to rule out the second candidate.

(6)   Korean perception of the English initial /d/, i.e. the sound [_ᵈō]

[_ᵈō]	*[ʰ] /−asp/	*[*no noise*] /+asp/	*[⁻] /+tense/	*['] /−tense/
☞ /to/				
/tʰo/		*(!)	*(!)	
/t'o/			*!	

The second candidate violates *[*no noise*]/+asp/, because the input sound [_ᵈó] contains no aspiration noise but the output structure /tʰo/ does contain the feature value /+asp/. The two aspiration cue constraints are powerless, however, in ruling out the third candidate; for that, we need a cue constraint that addresses the feature value /+tense/ which is present in the candidate structure /t'o/ (as well as in /tʰo/). This constraint is *[⁻]/+tense/, i.e. "an auditory normal F0 should not be

perceived as the feature value /+tense/". Since this is included in (6), /to/ remains as the only option for the perception of [_ᵈō].

To complete our set of cue constraints for initial plosives, we notice that the counterpart of the constraint *[⁻]/+tense/ is *[´]/−tense/, i.e. "an auditory raised F0 should not be perceived as the feature value /−tense/". We included this constraint vacuously in (6), but Tableau (7), an elaboration of Tableau (5), shows that it could play a role in the Korean perception of the English initial /t/.

(7) Korean perception of the English initial /t/, i.e. the sound [_ᵗʰó] (final version)

[_ᵗʰó]	*[ʰ] /−asp/	*[no noise] /+asp/	*[⁻] /+tense/	*[´] /−tense/
/to/	*(!)			*(!)
☞ /tʰo/				
/t̓o/	*!			

Together, Tableaus (6) and (7) illustrate that we can formulate the facts of perception alternatively in OT tableaus and in plain English. For instance, the first candidate row in (7) just states that two auditory cues contained in the sound [_ᵗʰó] (namely moderately strong noise and raised F0) militate against perceiving this sound as the phonological structure /to/ (which contains the feature values /−asp/ and /−tense/).

The constraint set in (6) and (7) is still a bit too coarse-grained. In real life, auditory events can take on continuous values along multi-dimensional auditory continua, so a full set of cue constraints needed to describe a language requires more auditory values than are displayed in the constraints of (6) and (7). For instance, we meant the constraint *[ʰ]/−asp/ to refer to an English-like aspiration noise of 80 ms (Lisker & Abramson 1964:394). However, stronger (longer) aspiration noises, i.e. [ʰː], are possible (in fact, they are typical of Korean /tʰ/: Lisker & Abramson 1964:397, Kagaya 1974:168) and will even be less likely to be perceived as /−asp/. In other words, the cue constraint *[ʰː]/−asp/ exists (and is ranked higher than *[ʰ]/−asp/: see §4.2). Working this out in full detail for the continua of aspiration noise and F0 is beyond the scope of this paper, whose focus is on vowel insertion. A more complete, 'principled', set of cue constraints than we could provide here appears in the next paragraphs, where we address the perception of the somewhat more straightforward auditory vowel height continuum.

In our discussion of the Korean perception of the English words *tag* and *deck*, we proceed with the English vowel sounds in these words, i.e. [ạ] and [ɛ̣]. An English listener interprets these as her phonemes /æ/ and /ɛ/, but how does a

Korean listener classify them? Korean has the vowels /i, ɨ, u, o, ɛ, æ, ʌ, a/, whose typical pronunciations are (or were) [i, ɨ, u, o, e̞, ɛ, ʌ, a̠] (based on Yang 1996).[5] The two most reasonable candidates for the perception of the two English nonhigh front vowels are the two Korean nonhigh front vowels /ɛ/ and /æ/. Which of the two does the Korean listener choose for [a̠], and which for [ɛ]?

This question can be answered in perception experiments, and has been answered as follows (Ingram & Park 1997): (older) naive Korean listeners of English perceive the (Australian) English sound [ɛ] (from English /ɛ/) as the Korean vowel /ɛ/ and the English sound [a̠] (from English /æ/) as the Korean vowel /æ/.[6]

The auditory continuum that is responsible for the auditory distinction between Korean /ɛ/ and /æ/ is vowel height; a full, 'principled', set of cue constraints has to link every possible auditory vowel height to each of the two phonological categories. For instance, the vowel /ɛ/ is linked to just as many vowel heights as the auditory nerve discretizes the vowel height continuum into. For reasons of space, we divide the vowel height continuum into ten steps only. The ten cue constraints for /ɛ/ are thus *[i]/ɛ/, *[ɨ]/ɛ/, *[e̞]/ɛ/, *[e]/ɛ/, *[e̞]/ɛ/, *[ɛ]/ɛ/, *[ɛ]/ɛ/, *[ɛ]/ɛ/, *[a̠]/ɛ/, and *[a]/ɛ/. In perception, the meaning of e.g. the constraint *[a̠]/ɛ/ is "the sound [a̠] should not be perceived as the vowel segment /ɛ/".

One may think that such large constraint sets are too powerful. That is, with so many cue constraints one could model *any* kind of perception. However, Boersma's (1997) proposal comes with a learning algorithm that ranks the cue constraints in such a way that the listener, after hearing a sufficiently large variety of tokens of every phonological category, becomes a *probability-matching listener*. That is, the listener will automatically rank her cue constraints in such a way that a given auditory event will be most likely perceived as the phonological category that was

---

5.   Irritatingly, the two vowels we are talking about in this section, namely /ɛ/ and /æ/, are nowadays in a state of impending merger (Yang 1996; Ingram & Park 1997; Lee & Ramsey 2000; Tsukada, Birdsong, Bialystok, Mack, Sung & Flege 2005). The pronunciations hypothesized in this section are meant to refer to the situation at the moment of the adaptation of the words *tag* and *deck*, i.e., we assume that /æ/ was pronounced as [ɛ], which is lower than the pronunciations measured by Yang, which can be transcribed as [ɛ] for males and [ɛ] for females.

6.   Tsukada et al. (2005:269) report quite different results for Korean listeners to an unspecified variety of (probably North-American) English, with English /æ/ mostly perceived as Korean /a/. Y. Kang's list of borrowings indeed show some cases of /æ/ borrowed as /a/. In order to understand what vowels are borrowed how, one would have to consider the English donor variety as well as the receiving Korean variety at the time of borrowing (see also §7.4 for a complicating factor). We speculate that a possible shift in the donor variety may be responsible for the different vowels in /.sɨ.pʰot./ 'spot', /.tʰɨ.lot./ 'trot' versus /.hat./ 'hot', /.sjat./ 'shot'.

most likely intended by the speaker (Boersma 1997:52–54; Escudero & Boersma 2003:79–81). For the case at hand, this means the following. The sound [ɛ�] is a possible realization of the Korean vowel /ɛ/ as well as of the Korean vowel /æ/. If Korean speakers, now, pronounce 70 [ɛ̞] tokens for /ɛ/ in the same time span as they pronounce 30 [ɛ̞] tokens for /æ/, a Korean listener-learner will come to perceive [ɛ̞] 70 percent of the time as /ɛ/, and 30 percent of the time as /æ/. That is, the learning algorithm will gradually rank the cue constraint *[ɛ̞]/æ/ above the cue constraint *[ɛ̞]/ɛ/. It will rank the complete set of cue constraints approximately as in Fig. 3.
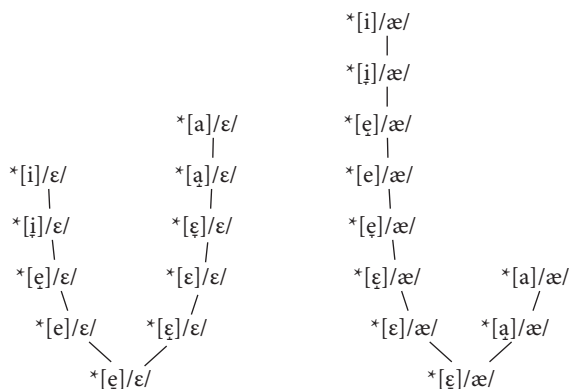
Figure 3. Rankings of the cue constraints that connect auditory vowel height to the Korean vowel categories /ɛ/ and /æ/

The figure assumes that the most typical realization of Korean /ɛ/ is [e̞], and the most typical realization of Korean /æ/ is [ɛ̞]. As a result, the constraints *[e̞]/ɛ/ and *[ɛ̞]/æ/ get ranked lowest. The remaining constraints get ranked by confusability and frequency (Boersma 2006; Boersma & Hamann 2008), which basically entails that they indirectly get ranked by auditory distance; thus, *[i]/æ/ will be ranked very high, because speakers are very unlikely to pronounce an intended /æ/ as the sound [i].

Tableaus (8) and (9) show that with the rankings of Fig. 3, the English sound [ɛ̞] is perceived as /ɛ/ and the English sound [a̞] as /æ/.

(8)   Korean perception of the English vowel /ɛ/, i.e. the sound [ɛ̞]

[ɛ̞]	*[a̞]/ɛ/	*[ɛ]/ɛ/	*[e̞]/ɛ/	*[ɛ]/ɛ/	*[ɛ̞]/ɛ/	*[a]/æ/	*[e]/ɛ/	*[ɛ]/ɛ/	*[ɛ]/æ/	*[a̞]/æ/	*[e̞]/ɛ/	*[ɛ]/æ/
☞ /ɛ/								*					
/æ/				*!									

(9)   Korean perception of the English vowel /æ/, i.e. the sound [ạ]

[ạ]	*[ạ] /ɛ/	*[ɛ̧] /ɛ/	*[ẹ] /ɛ/	*[ɛ] /ɛ/	*[ɛ] /æ/	*[a] /æ/	*[e] /ɛ/	*[ɛ] /ɛ/	*[ɛ] /æ/	*[ạ] /æ/	*[ẹ] /ɛ/	*[ɛ] /æ/
/ɛ/	*!											
☞ /æ/										*		

We have thus formalized perception on the basis of 'least confusable', not on the basis of 'most similar', or 'auditorily nearest'. This contrasts with approaches that assume that speakers have direct knowledge of the auditory distance between phonological elements, such as Steriade's (2001) P-map or Flemming's (1995) MINDIST constraints (for discussion see Boersma & Hamann, 2008:§7.4).

We now turn to the final consonants of English *tag* and *deck*. The cues in the final consonants are a superset of those of the initial consonants. In [_ᵗʰǽːɡ˺‿ᵍ], a Korean listener has no longer only the lenis burst cue [ᵍ], but also: (1) the closure voicing [‿], which is compatible with the Korean /k/, which is the only of the three plosives that can ever be voiced (see §3.2); and (2) the vowel lengthening [ː], which occurs in Korean only before lax phonemes such as /k/ (see §4.3). In [_ᵈɛ̠k˺_ᵏʰ], a Korean listener has the fortis burst [ᵏ] and the aspiration [ʰ], which are the same cues as for /kʰ/ in initial position. So it might seem reasonable that [_ᵗʰǽːɡ˺‿ᵍ] and [_ᵈɛ̠k˺_ᵏʰ] are perceived as /.tʰæk./ and /.tɛkʰ./, respectively. This is indeed a view that is widely held in theories on loanword adaptation (Silverman 1992; Yip 1993; H. Kang 1996; Yip 2006). With Y. Kang (2003), however, we regard it as unlikely. The next section explains why.

3.2   Korean perception of word-final release bursts: Vowel hallucination

In §3.1 we asserted that the listener's perception process is defined as an attempt to retrieve the speaker's intended surface form. If this is correct, the Korean interpretations of the English final sound sequences [ːɡ˺‿ᵍ] and [k˺_ᵏʰ] are unlikely to be just the segments /k/ and /kʰ/. This is because it is very unlikely that the sound sequences [ːɡ˺‿ᵍ] and [k˺_ᵏʰ] can represent an intended *Korean* final /k/ and /kʰ/, because release bursts such as [ᵍ] and [ᵏ] do not occur in Korean codas.

Korean final plosives are pronounced without a release burst (Martin 1951; H. Kim 1998; Y. Kang 2003). Thus, the form /.pat./ in (1) has the auditory-phonetic form [_ᵇat˺_] (where [t˺] stands for the formant transition from the vowel into the coronal closure), not the fully released *[_ᵇat˺_ᵗ]. For the listener, therefore, the presence of a release burst in Korean always indicates that the consonant is an onset and that it is followed by a vowel. We can express this fact as the cue

constraint *[*burst*]/C(.)/, which stands for "an auditory release burst should not be perceived as a phonological consonant in coda."[7]

To satisfy the strong constraint *[*burst*]/C(.)/, the Korean listener has the option to perceive an onset instead of a coda. This entails perceiving [_thǽːɡ$^{\urcorner}$˳^{g}] and [_^{d}ɛ�posk$^{\urcorner}$_^{kh}] as /.tʰæ.kɨ./ and /.tɛ.kʰɨ./, respectively. Both perceptions violate a cue constraint against interpreting nothingness as a vowel: *[ ]/ɨ/. To assess how highly ranked such a constraint could be, we have to realize that background noise often obliterates auditory cues in speech. For instance, the hypothetical Korean phonological sequences /.o.kɨ./ and /.o.kʰɨ./ will ideally be produced as [oɡ$^{\urcorner}$˳^{g}ɨ] and [ok$^{\urcorner}$_^{kh}ɨ], but may sometimes sound like the impoverished [oɡ$^{\urcorner}$˳^{g}] and [ok$^{\urcorner}$_^{kh}], especially across a larger distance or if there is background noise. Such losses of direct positive auditory information are likely to occur in every language, and in Korean this is especially likely to happen if the final vowel is /ɨ/, which has been reported to be 'often deleted, especially in a weak, non-initial open syllable' (Kim-Renaud 1987, as quoted by Y. Kang 2003:236). The learning algorithm discussed in §3.1 will then rank *[ ]/ɨ/ low. As a result, listeners will routinely fill in the missing information.

The interpretations of [_thǽːɡ$^{\urcorner}$˳^{g}] as /.tʰæ.kɨ./ and of [_^{d}ɛ̠k$^{\urcorner}$_^{kh}] as /.tɛ.kʰɨ./ could now be described in terms of the same cue constraints as in (6) and (7), with the addition of *[*burst*]/C(.)/ and *[ ]/ɨ/. However, we must realize that if a vowel is perceptually epenthesized, the final consonant becomes phonologically intervocalic, and this has consequences for the cues because in phonologically intervocalic position the Korean lax plosive is voiced (Kagaya 1974; Iverson 1983; Y.Y. Cho 1990; Jun 1995). Moreover, in noninitial syllables F0 cues are reduced (M.-R. Kim 2000; Kim & Duanmu 2004). The cue constraints that relate tenseness to F0 in (6) and (7) must therefore be reformulated as *[ˉ]/([$_\phi$)+tense/ and *[ˊ]/([$_\phi$)−tense/ (where "[$_\phi$" denotes a phonological phrase boundary), and for the voicing cue we need the cue constraint *[*no voice*]/(V)−tense(V)/, which states that a voiceless silence cannot be perceived as a lax plosive between two phonologically present vowels, and its counterparts *[˷]/C(.)/ and *[˷]/+tense/, which state that a voiced closure cannot be perceived as a coda consonant and cannot be perceived as a fortis or aspirated plosive. The formalization is given in Tableaus (10) and (11), which do not contain the cue constraints that refer to F0 as they are irrelevant for these cases.

---

7.   In the formulation of this constraint, the parentheses denote the environment; the remaining two elements, i.e. C and *burst*, are in *correspondence*, in the sense of Correspondence Theory (McCarthy & Prince 1995). An alternative formulation of the constraint is therefore *[*burst$_i$*]/C$_i$ ./.

(10)   Korean perception of the English word *tag*

[_ᵗʰág˺_g]	*[burst] /C(.)/	*[ʰ] /−asp/	*[no noise] /+asp/	*[͜] /C(.)/	*[͜] /+tense/	*[no voice] /(V)−tense(V)/	*[ ] /ɨ/
/.tʰæk./	*!			*			
/.tʰækʰ./	*!		*	*	*		
☞ /.tʰæ.kɨ./							*
/.tʰæ.kʰɨ./			*(!)		*(!)		*
/.tʰæ.k'ɨ./					*!		*

(11)   Korean perception of the English word *deck*

[_ᵈɛk˺_kʰ]	*[burst] /C(.)/	*[ʰ] /−asp/	*[no noise] /+asp/	*[͜] /C(.)/	*[͜] /+tense/	*[no voice] /(V)−tense(V)/	*[ ] /ɨ/
/.tɛk./	*!	*					
/.tɛkʰ./	*!						
/.tɛ.kɨ./		*(!)				*(!)	*
☞ /.tɛ.kʰɨ./							*
/.tɛ.k'ɨ./		*!					*

In (10) and (11), the new cue constraint *[*burst*]/C(.)/ rules out the plosive-final candidates. The cue constraints *[*no noise*]/+asp/ and *[͜]/+tense/ rule out the remaining candidates with aspirated and fortis plosives in (10), and *[ʰ]/−asp/ and *[*no voice*]/(V)−tense(V)/ rule out the remaining candidates with the unaspirated plosives in (11). The cue constraint *[ ]/ɨ/ asserts that one should not hallucinate the vowel /ɨ/ if there is no direct corresponding auditory cue. It is the weakness of this constraint that causes the insertion of 'illusory' vowels in perception.

3.3   Korean loanword adaptation: Structural constraints

If perception could be handled by cue constraints alone, perception would hardly interact with the phonology. That is, the surface elements that appear in the formulations of the cue constraints are phonological elements, but that would be all. However, according to Fig. 1 the integration of perception and phonology is much stronger than that: the output of the perception process itself is evaluated by *structural constraints*. As argued by Polivanov (1931), Boersma (2000, 2007ab), and Pater (2004), the same structural constraints that restrict phonological production (the top right of Fig. 1) also restrict prelexical perception (the bottom left of Fig. 1). That is, perception is not handled by cue constraints alone, but by an interaction between structural and cue constraints. This renders perception

thoroughly phonological itself. In other words, there is no longer any distinction between perception and phonology. In fact, the often discussed question whether loanword adaptation is 'due to the phonology or due to perception' is rendered moot (see also §6). The present section illustrates how structural constraints play a role in the perceptual vowel insertion in Korean loanwords from English.

Several structural constraints have been introduced in the phonological production tableaus of §2, but none of them were used in the perception tableaus of §3.1 and §3.2. One structural constraint could already have made its appearance in Tableau (11), namely the constraint */+asp ./ that was crucial in Tableau (1). If included in Tableau (11), it would have helped to rule out the candidate /.tɛkʰ./. But of course, this constraint would not have played a crucial role in that tableau, which works perfectly with cue constraints alone.

More crucial cases of structural constraints that guide perception were given by Polivanov (1931) in a discussion of Japanese perception of Russian ([ˍᵗak˺ˍᵏ] → /.ta.ku./, [ˍᵈrama] → /.do.ɾa.ma./), a case that was translated to the OT perception framework of Fig. 1 by Boersma (2007b).

A similar case as the Japanese vowel insertion in consonant clusters that Polivanov analysed, is found in the Korean avoidance of complex onsets in both native and loanwords. Thus, the English word *spike* is realized in Korean as /.sɨ.pʰa.i.kʰɨ./, and *flute* is realized as /.pʰɨl.lu.tʰɨ./ (Y. Kang 2003:262,266,244). Since this insertion generalizes to all onset clusters, the most straightforward way to account for it is by utilizing the structural constraint */.CC/.[8] Tableau (12) shows the analysis for *spike*, where we formalize only the adaptation of the initial cluster and thereby ignore the adaptation of the diphthong and the final consonant.

(12)　Korean perception of the English word *spike*

[s̯ᴾaɪk˺ˍkh]	*/.CC/	*[ ] /ɨ/
/.spʰa.i.kʰɨ./	*!	
☞　/.sɨ.pʰa.i.kʰɨ./		*

In (12) we see that the structural constraint, by outranking the cue constraint, causes the perceptual insertion of an illusory vowel.[9]

8.　It is always possible, though often awkward, to replace a structural constraint with a set of cue constraints. We elaborate on this possibility in §7.1.

9.　The attentive reader may notice that it is strange that the sound [ˍᴾ] is perceived as /pʰ/, which is typically pronounced [ˍᵖʰ˙], rather than as /p'/, which is typically pronounced [ˍᴾ] (§3.1). This problem is discussed by Oh (1996), H.Kang (1996), Kenstowicz (2005), Ito, Kang & Kenstowicz (2006), Davis & Cho (2006), and Iverson & Lee (2006), and we return to it in §7.4.

To show that the same structural constraints play a role in perception and production, we consider the Korean avoidance of strident codas, which was illustrated for production in §2.1. The same structural constraint that caused coda neutralization in production (Tableau (2)), causes vowel insertion in perception, as illustrated in Tableau (13), which shows the Korean perception of the English word *mass*.

(13)   Korean perception of the English word *mass*

[mæs]	*/+stri ./	*[*friction*] /−stri/	*[ ] /ɨ/
/.mæs./	*!		
/.mæt./		*!	
☞ /.mæ.sɨ./			*

In Tableau (13), the most 'faithful' percept /.mæs./ violates the structural constraint */+stri ./. The candidate /.mæt./ violates the high-ranked cue constraint *[*friction*]/−stri/ which says that friction noise should not be interpreted as phonological nonstridency. The winning candidate is therefore the percept /.mæ.sɨ./, with an epenthesized vowel.[10]

A comparison between the perception tableau in (13) and the production tableau in (2) shows us that the forbidden strident coda consonant /-s./ is ruled out in both tableaus by the high-ranked structural constraint */+stri ./. The repair mechanisms, however, are different in L1 perception and L1 production. In both perception and production, the choice goes between the remaining surface forms /-t./ and /-.sɨ./. In perception, the constraint for honouring phonetic stridency information (*[*friction*]/−stri/) outranks the constraint against vowel insertion (*[ ]/ɨ/), leading to the surface form /-.sɨ./, whereas in production, the constraint against vowel insertion (Dᴇᴘ-V) outranks the constraint for honouring underlying phonological stridency (Iᴅᴇɴᴛ(stri)), leading to the surface form /-t./. Please note that these differences are not due to different constraint rankings between comprehension and production, but to different kinds of constraints in the 'phonological' part of the grammar (the top of Fig. 1) and the 'phonetic' part of the grammar (the bottom of Fig. 1). Please also note that this does not mean that the 'phonological' and 'phonetic' parts of the grammar can be viewed as separate modules: they utilize the same structural constraints.

---

10.   When comparing (13) with (12), we see that English [s] is adapted into Korean as plain /s/ if followed by a stop (in English), but as tense /sʼ/ if it is final (in English). The present paper makes no attempt to account for this difference. See Davis & Cho (2006) and H. Kim (this volume) for more information.

Another potential case of a structural constraint in perception is the case of the constraint S�YʟʟCᴏɴ, which bans segmental sequences like /km/ from the output of Korean L1 production (§2.3). H. Kang (1996) notes that English words with word-internal plosive-nasal clusters are borrowed differently in Korean (namely, with vowel insertion) than English words with plosive-plosive clusters (which are borrowed without vowel insertion). As usual, we interpret this as the result of a difference in perception. Thus, we follow Y. Kang (2003) in assuming that *chapter* is perceived as /.tsʰæp.tʰʌ./; we also assume that, by contrast, the word *picnic* (which Y. Kang mentions but does not analyse) is perceived as /.pʰi.kʰɨ.nik./,[11] as was confirmed in the lab by Hwang et al. (2007).

According to H. Kang (1996), the adaptation of *chapter* as /.tsʰæp.tʰʌ./ with a coda /p/ is due to the fact that the English source word is pronounced without an audible labial release, i.e. as [_ᵗʃap̚_ᵗʰə]. We follow Y. Kang (2003) in assuming that this lack of release causes Korean listeners to perceive a coda /p/. We formalize this in Tableau (14). With the same cue constraint that caused the insertion of a vowel in (10) and (11), namely *[*burst*]/C(.)/, the winning candidate now becomes the form without vowel insertion.

(14)   Korean perception of the English word *chapter*

[_ᵗʃap̚_ᵗʰə]	S�YʟʟCᴏɴ	*[C̚] / /	*[*burst*] /C(.)/	*[ ] /ɨ/
☞ /.tsʰæp.tʰʌ./				
/.tsʰæ.pɨ.tʰʌ./				*!
/.tsʰæ.pʰɨ.tʰʌ./				*!
/.tsʰæ.tʰʌ./		*!		

In contradistinction with Tableaus (10) and (11), the candidate without vowel insertion (/.tsʰæp.tʰʌ./) now wins: since the input sound contains no labial release burst, this candidate does not violate *[*burst*]/C(.)/. The fourth candidate violates the cue constraint *[C̚]/ /, which is high-ranked because Korean listeners routinely have to interpret postvocalic formant transitions as true consonants.

The same ranking as in (14) works out differently for an English plosive-nasal cluster, such as in *picnic*. This word is pronounced in English as [_ᵖʰɪk̚_nɪk̚_], where we assume the same lack of release for the first consonant of the cluster (as a side issue, we also assume with Y. Kang 2003:261 that the word-final plosive is unreleased in English, so does not violate *[*burst*]/C(.)/ either). The difference with

---

11.   H. Kang notes that younger generations can produce this word as /.pʰiŋ.nik./. According to Kabak (2003:59), such adaptations are due to orthographical influence. See §7.3 for how this fits into our model.

the *chapter* case is that the syllable contact law is applicable. Tableau (15) shows that this forces vowel insertion.

(15)   Korean perception of the English word *picnic*

[_ᵖʰɪk˥_nɪk˥_]	SYLLCON	*[C˥] //	*[burst] /C(.)/	*[_] /+nas/	*[no noise] /+asp/	*[no voice] /(V)_tense(V)/	*[ ] /ɨ/
/.pʰik.nik./	*!						
/.pʰiŋ.nik./				*!			
☞ /.pʰi.kɨ.nik./						*	*
☞ /.pʰi.kʰɨ.nik./					*		*
/.pʰi.nik./			*!				

A new type of candidate in Tableau (15) is /.pʰiŋ.nik./. If a Korean listener interpreted [_ᵖʰɪk˥_nɪk˥_] as /.pʰiŋ.nik./, she would ignore positive information, namely the silence [_], which is a reliable cue for the presence of a plosive rather than a nasal. In (15) this is expressed with the cue constraint *[_]/+nas/. The thing that crucially distinguishes Tableau (15) from Tableau (14), though, lies somewhere else. The crucial difference is that the candidate /.pʰik.nik./ violates the syllable contact law. This causes the listener to insert a vowel. The choice between the third and fourth candidate in (15) has to be made on the basis of lower-ranked cue constraints; the attested /.pʰi.kʰɨ.nik./ suggests that the absence of auditory voicing weighs heavier than the absence of auditory aspiration noise.[12]

    It is crucial for our story that the difference between (14) and (15) can only be accounted for in terms of the structural constraint SYLLCON. This is because the auditory cues for the first cluster consonant are the same in (14) and (15), namely place-lending formant movements followed by a silence; hence, cue constraints alone cannot explain why a vowel is perceived in (15) but not in (14), and if we reranked the cue constraints in such a way that they alone could cause vowel insertion in *picnic*, they would also cause vowel insertion in *chapter*. The bottom line is that cue constraints do not suffice, and that therefore structural constraints are crucial in perception.

    When comparing (4) with (15), we see that forbidden sonority sequences like /kn/ are repaired differently in L1 production than in L1 perception. Both in production and perception, SYLLCON rules out candidates with sequences like /kn/. In production, however, the ranking of the faithfulness constraints DEP-V >> IDENT(nas) decides that the repair is /ŋn/, whereas in perception the ranking of the cue constraints *[_]/+nas/ >> *[ ]/ɨ/ decides that the repair is /kʰɨn/.

---

12.   The possible candidate /.pʰi.k'ɨ.nik./ has to be ruled out in a different way. See §7.4.

Thus, structural constraints are crucial in perception, and the repair strategies can be different in perception and production.

3.4   What is perception?

Not all readers will instantly accept our view (shared by Y. Kang 2003, Kabak & Idsardi 2007 and H. Kim this volume) that perception can introduce a vowel, as in (10), (11), (12), (13) and (15).

However, precisely such perceptual vowel insertion has been proposed several times before. Polivanov (1931) argues that Japanese listeners perceive the Russian word [ˌᵗakˤ⁼ᵏ] 'so' as their native phonological structure /.ta.ku./, and the Russian word [ˌᵈrama] 'drama' as their structure /.do.ɾa.ma./. Polivanov attributes these perceptions to Japanese structural constraints against coda consonants and against complex clusters, respectively; indeed, a formulation in terms of an interaction between structural and cue constraints in OT, analogous to Tableaus (12) and (15), is possible and has been carried out in detail by Boersma (2007b:10–14).

Polivanov's proposal has been confirmed in the laboratory. Dupoux, Kakehi, Hirose, Pallier & Mehler (1999) showed that Japanese listeners could not discriminate between the sounds [ebzo] and [ebuzo], which strongly suggests that Japanese listeners perceive the sound [ebzo] as their native phonological surface structure /.e.bɯ.zo./.

We would like to stress here, however, that linguistic perception is not about discrimination, but about identification (for Korean vowel insertion, Kabak & Idsardi 2007:36 agree with this view). We regard perception as an active process: generally, perception is the mapping from raw sensory data to a more abstract mental representation that is ecologically appropriate; in linguistics, the listener's active perception process maps a sound to a native phonological structure, in order to arrive quickly at the morphemes that the speaker has intended to bring across. When computing a likely intended phonological structure, the listener has to take into account both the available auditory cues and knowledge about the structural restrictions of the language. With Boersma (2000, 2007ab), therefore, we formalize this computation in terms of interactions between structural and cue constraints, as in Fig. 1 and Tableaus (12) to (15). Peperkamp & Dupoux (2003) propose the same three levels and four mappings for loanword adaptations as we employ in Fig. 1, noting that such representations and mappings correspond to what psycholinguists would have to say about the stages of comprehension (McQueen & Cutler 1997) and production (Levelt 1989); however, they do not provide a linguistic modelling of these mappings and in fact regard perception as nonlinguistic.

In Optimality Theory, the idea that structural constraints play a role in perception has some history. It is related to the idea of *robust interpretive parsing* (Tesar 1997; Tesar & Smolensky 2000), in which listeners interpret an *overt form* (sound)

as a phonologial (e.g. metrical) structure by using the same ranking of the structural constraints as they use for production.[13] Cue constraints turned up in Boersma (1997, 1998, 2000, 2006, 2008), Escudero & Boersma (2003, 2004), Escudero (2005), Boersma & Escudero (2008), Boersma & Hamann (2008), and Hamann (2009), and their interaction with structural constraints was formalized in various degrees of similarity to the present proposal by Boersma (1998:164–171,364–396, 2000, 2007ab), Boersma, Escudero & Hayes (2003), and Pater (2004).

We like to urge phonologists to regard active perception as a just as intricate and interesting process as they have traditionally regarded production. A spectacular example was given by Boersma (2000:21–22, 2007b:27), who argues that Desano listeners interpret the sound [ʒuẽũ] (the Portuguese name *João*) as their native surface structure /.ɲũ./, forced by a structural constraint against tautosyllabic sequences of oral and nasal segments. In the loanword adaptation literature, perception is often regarded as a much less active, and therefore much less powerful, process. This view of perception has led researchers to fail to consider L1-specific perception phenomena as the explanans for loanword adaptation. In §6 we compare our account to some of these other proposals, as well as to proposals that do accept vowel insertion in perception but do not formalize it.

3.5   Conclusion

Section 3 has illustrated how Korean listeners interpret English sound sequences in terms of their own phonological cues and phonotactics. To illustrate that the exact same constraints and ranking work for the perception of native Korean words as well, Tableau (16), which uses the same ranking as (10) and (11), shows that the auditory-phonetic form [mɛk˺_] (the normal unreleased pronunciation of the Korean word |mæk| 'pulse') is perceived as /.mæk./.

(16)   L1 Korean perception

[mɛk˺_]	*[burst] /C(.)/	*[ʰ] /−asp/	*[no noise] /+asp/	*[_] /+tense/	*[no voice] /(V)−tense(V)/	*[ ] /ɨ/
☞ /.mæk./						
/.mækʰ./			*!			
/.mæ.kɨ./					*!	*
/.mæ.kʰɨ./			*!			*
/.mæ.kʼɨ./						*!

---

13.   The relation between robust interpretive parsing and perception is discussed in Boersma (2007b:21–23).

While for the Korean perception of English forms in (10) and (11) the ranking resulted in vowel insertion, for the perception of a native Korean form in (16) it does not result in vowel insertion. The reason for this difference simply lies in the auditory input: Korean final plosives are unreleased, whereas final plosives in English are released (but see §4.3 for unreleased plosives in English).

In the next section we show that the perceptual adaptations of English words are sufficient to explain vowel insertion in loanwords.

4. Perception, storage and production of English loanwords in Korean

In §3 we illustrated the very first part of the loanword adaptation process, namely the mapping from an auditory-phonetic form (sound) to a native phonological surface structure. Loanword adaptation does not stop here: this foreign-language perception has to be followed by a process of lexical storage, which can then lead to the adapter's own productions of the borrowed word. This is the same process that any listener uses for the words of her native language.

4.1 Storage of English loanwords in the Korean lexicon

We assume that loanword adaptation has started with the L1 perception process exemplified in §3. For instance, the Korean loanword adapter has perceived [_tʰɛ́ːg˥ ˺g] and [_dɛ̰k˥ ˺kh] as /.tʰæ.kɨ./ and /.tɛ.kʰɨ./ (§3.2). Her next task is to store them in her lexicon as new underlying forms.

We assume that the storage of a new word in the lexicon follows the process that we call *recognition* in the top left of Fig. 1. In this process, the faithfulness constraints ensure that the learner stores into her lexicon the fully faithful forms |tʰækɨ| and |tɛkʰɨ|, as illustrated in Tableaus (17) and (18).[14]

(17)   Korean lexical storage of the English word *tag*

/.tʰæ.kɨ./	*/+asp ./	Dep-V	Max-C	Ident(asp)	*/C ./
|tʰækʰɨ|				*!	
☞ |tʰækɨ|					
|tʰæk|		*!			

---

14.   Full faithfulness in word recognition is ensured only if the faithfulness constraints do not conflict with constraints at higher levels, which would come into play if alternations start to play a role. See §5 for examples.

(18)   Korean lexical storage of the English word *deck*

/.tɛ.kʰɨ./	*/+asp ./	Dᴇᴘ-V	Mᴀx-C	Iᴅᴇɴᴛ(asp)	*/C ./
☞ |tɛkʰɨ|					
|tɛkɨ|				*!	
|tɛkʰ|	*!				
|tɛk|	*!			*	

The structural and faithfulness constraints are the same as in (1), and they are ranked in the same order. We first note that the third candidate in (18) does not violate */+asp ./, because this constraint only evaluates surface forms; something analogous holds for the constraint */C ./, which incurs no violations for any candidates. Next, we see that Dᴇᴘ-V can still be high-ranked (note that in this direction of processing, Dᴇᴘ-V militates against *deletion* rather than insertion: correspondence constraints evaluate relations, not processes).

The remaining example words from §3 are stored fully faithfully as well: |sɨpʰaikʰɨ|, |mæsʼɨ|, |tsʰæptʰʌ|, |pʰikʰɨnik|.

Building a lexicon mainly through faithfulness constraints, as in (17) and (18), constitutes a form of *lexicon optimization* (Prince & Smolensky 1993 [2004:225–231]). As a result, the lexicon comes to reflect some of the same phonotactic restrictions that surface forms have, an effect that Boersma (1998:395) called *poverty of the base* (for exceptions see §7.3).

4.2   Production of English loanwords from a Korean lexicon

After storing the English word as the new underlying forms |tʰækɨ| and |tɛkʰɨ|, the loanword adapter is ready to subsequently use them in her own productions. She will produce them as the surface forms /.tʰæ.kɨ./ and /.tɛ.kʰɨ./ and as the auditory-phonetic forms [_ᵗʰɛ́ːɡ⌐‿ɡɨ] and [_ᵈɛ̀k⌐_ᵏʰːɨ], as the following four tableaus illustrate.

In phonological production, the underlying |tʰækɨ| and |tɛkʰɨ| are produced as /.tʰæ.kɨ./ and /.tɛ.kʰɨ./, as Tableaus (19) and (20) show.

(19)   Korean phonological production of the English loanword tʰækɨ

|tʰækɨ|	*/+asp ./	Dᴇᴘ-V	Mᴀx-C	Iᴅᴇɴᴛ(asp)	*/C ./
☞ /.tʰæ.kɨ./					
/.tʰæ.kʰɨ./				*!	
/.tʰæk./					*!
/.tʰæ./			*!		

(20) Korean phonological production of the English loanword tɛkʰɨ

|tɛkʰɨ|	*/+asp ./	Dᴇᴘ-V	Mᴀx-C	Iᴅᴇɴᴛ(asp)	*/C ./
/.tɛ.kɨ./				*!	
☞ /.tɛ.kʰɨ./					
/.tɛkʰ./	*!				*
/.tɛk./				*(!)	*(!)

In these production tableaus, we employ the same constraints as in the production tableau for the native form |patʰ| in (1), and in the recognition tableaus (17) and (18). Deletion of final underlying |ɨ| is prevented by the low-ranked */C ./ (this obviates the need for Mᴀx-V, at least for the cases discussed here). Likewise, the other sample words are produced equally faithfully as /.sɨ.pʰa.i.kʰɨ./, /.mæ.s'ɨ./, /.tsʰæp.tʰʌ./, /.pʰi.kʰɨ.nik./.

The surface form /.tʰæ.kɨ./ that results from the phonological production in (19) is subsequently pronounced as [_ᵗʰːɛ́ːg⌐ ̰gɨ], as Tableau (21) shows. Here we employ the same cue constraints as in perception, i.e. cue constraints are just as bidirectional as the faithfulness constraints and the structural constraints. For instance, the constraint *[h]/−asp/, which meant in perception "if there is moderately strong auditory-phonetic noise, then do not perceive the phonological surface structure /−asp/", now means in production "the phonological surface structure /−asp/ should not be realized with moderately strong auditory-phonetic noise". This constraint is ranked at the same height in production and perception.

(21) Korean phonetic implementation of the English loanword tʰækɨ

/.tʰæ.kɨ./	*[hː] /−asp/	*[burst] /C(.)/	*[h] /−asp/	*[no noise] /+asp/	*[no voice] /(V)−tense(V)/	*[ ] /ɨ/	*[h] /+asp/	*[hː] /+asp/
[_ᵗʰːɛ́k⌐ ᵏʰːɨ]	*!				*			*
[_ᵗʰːɛ́k⌐ ᵏʰɨ]			*(!)		*(!)			*
[_ᵗʰːɛ́k⌐ ᵏɨ]					*!			*
☞ [_ᵗʰːɛ́ːg⌐ ̰gɨ]								*
[_ᵗʰːɛ́ːg⌐ ̰g]						*!		*
[_ᵗʰːɛ́k⌐ ᵏ]					*!	*		*
[_ᵗʰːɛ́k⌐]					*!	*		*

The *phonetic implementation tableau* (21) employs some of the same cue constraints as the perception tableaus (10) and (11), ranked in the same order. However, the cue constraints in (10) and (11) only had to deal with English

sounds. Here, in order to make the correct Korean-specific choice of auditory forms, we need cue constraints that cover the whole spectrum of auditory values. For aspiration, we have the ranking *[ʰː]/−asp/ >> *[ʰ]/−asp/, because it is worse to aspirate an unaspirated consonant strongly than to aspirate it only moderately (the two new /+asp/ cue constraints are explained below). Intervocalic voicing of the lax plosive in production is here achieved by the same *[*no voice*]/(V)−tense(V)/ that works in the perception tableaus (11), (15) and (16); it is violated by all voiceless candidates, even the ones with phonetic vowel deletion (because intervocality is defined at the phonological level). Please note that none of the candidates violate *[*burst*]/C(.)/, because the input is not consonant-final. Further, phonetic /ɨ/-deletion is punished by the constraint *[ ]/ɨ/ that we saw before; in perception, the low ranking of this constraint allowed the perception of an illusory vowel (§3.2); here in phonetic implementation, this constraint suddenly becomes crucial in making sure that the surface vowel /ɨ/ is pronounced at all. The second best candidate is [_ᵗʰ:ɛ̀ːg˺_ᵍ]; its second-bestship expresses the idea that the voicing cue (between *phonologically present* vowels) is more important than the audibility of the vowel /ɨ/; this candidate could be realized if an articulatory (laziness) constraint (Kirchner 1998; Boersma 1998) is ranked at about the same height as *[ ]/ɨ/ (in the model of Fig. 1, articulatory constraints such as *[ɨ] evaluate the phonetic form directly).

   The surface form /.tɛ.kʰɨ./ that results from (20) is pronounced as [_ᵈɛ̀k˺_ᵏʰːɨ], as (22) shows.

(22)   Korean phonetic implementation of the English loanword tɛkʰɨ

/.tɛ.kʰɨ./	*[ʰː] /−asp/	*[burst] /C(.)/	*[ʰ] /−asp/	*[no noise] /+asp/	*[_] /+tense/	*[ ] /ɨ/	*[ʰ] /+asp/	*[ʰː] /+asp/
☞ [_ᵈɛ̀k˺_ᵏʰːɨ]								*
[_ᵈɛ̀k˺_ᵏʰɨ]							*!	
[_ᵈɛ̀k˺_ᵏɨ]				*!				
[_ᵈɛ̀ːg˺_ᵍɨ]				*(!)	*(!)			
[_ᵈɛ̀k˺_ᵏʰː]						*!		*
[_ᵈɛ̀k˺_ᵏ]				*!		*		
[_ᵈɛ̀k˺_]				*!		*		

The full cue constraint ranking that we need for aspiration is *[*no noise*]/+asp/ >> *[ʰ]/+asp/ >> *[ʰː]/+asp/: the less noise there is in the auditory form, the less likely it is that a /+asp/ segment is present. The end result is that Koreans pronounce the borrowed English word *deck* with more aspiration than was present in the English

source sound, i.e., while perception has introduced an illusory vowel in this word, the phonetic implementation has given the word an additional Korean accent.

Together, Tableaus (19) to (22) perform production as a serial process: first phonological production, then phonetic implementation, where the output of the former is the input to the latter. The two partial production processes can also be done in parallel (Boersma 2007ab, 2008), with the same ultimate result as Tableaus (19) to (22) yield. An example of a parallel production tableau is given in §7.1.

Tableaus (21) and (22) round up our account of the Korean adaptation of the English words *tag* and *deck*. It started with the observed English sounds [_thá:g⌐ ͜g] and [_^{d}ɛ� k⌐_^{kh}], and ended with the observed Korean sounds [_$^{th:}$ɛ́:g⌐ ͜gɨ] and [_^{d}ɛ̀k⌐_$^{kh:}$ɨ]. What happens in between these two pairs of sounds is something that is not directly observable, and therefore open to widespread hypothesization by linguists. The account that we have brought here is the only one that utilizes exclusively first-language processing, as summarized in (23).

(23)   Korean adaptation of the English words *deck* and *tag*
       tag: [_thá:g⌐ ͜g] → /.tʰæ.kɨ./ → |tʰækɨ| → /.tʰæ.kɨ./ → [_$^{th:}$ɛ́:g⌐ ͜gɨ]
       deck: [_^{d}ɛ̝k⌐_^{kh}] → /.tɛ.kʰɨ./ → |tɛkʰɨ| → /.tɛ.kʰɨ./ → [_^{d}ɛ̀k⌐_$^{kh:}$ɨ]

The first two steps in (23) form part of the comprehension process, and the last two form part of the production process. The four steps of loanword adaptation are handled by a single constraint ranking, which is the same as the one used for native Korean production and comprehension. In this model, auditory similarity between the phonetic forms of the donor language and the borrowing language is achieved by the bidirectionality of the cue constraints; this bidirectionality obviates the need for supposed mechanisms by which speakers have direct knowledge of the auditory distance between phonological elements, such as Steriade's (2001) "P-map", which has been invoked very often in loanword adaptation research (§6.1).

A complete ranking is given in §7.5.

4.3   Variation

For English words with a final /g/ or /k/, a vowel is not always appended in Korean. Kim-Renaud (1977:252) and Y. Kang (2003:235) attribute this variation to the variability of the release burst in English (Rositzke 1943; Crystal & House 1988; Byrd 1992; Cruttenden 1994:145; H. Kim 1998). To understand this, we investigate how Tableaus (10) and (11) will change if the release burst is inaudible. First, the constraint *[*burst*]/C(.)/ will not be violated in any candidate. But there is more. One first has to realize that articulatorily, a release must exist, even if it is inaudible (H. Kim 1998). The release burst, then, is rendered inaudible by a low subglottal pressure. In the input of Tableau (11), this low pressure must have an influence on

the following aspiration noise, which will become inaudible itself: in (11), therefore, the auditory input will be [_ᵈɛ̠k˺_], the constraint *[ʰ]/−asp/ can no longer be violated, and the candidate /.tɛk./ will win (because /.tɛkʰ./ is ruled out by the high-ranking */+asp ./). In the input of Tableau (10) the low subglottal pressure during the release will reduce the subglottal pressure during the closure phase as well, so that closure voicing diminishes: in (10), therefore, the auditory input will be [_ᵗʰæ̆ːg˺ ̆] (where the breve stands for reduction), the constraints *[̬]/C(.)/ and *[̬]/+tense/ can no longer be violated (although the lower-ranked *[̆]/C(.)/ and *[̆]/+tense/ can), and the candidate /.tʰæk./ will win. For both cases we will then end up with final unaspirated plosives in the underlying forms, and subsequently with unreleased plosives in the produced auditory-phonetic forms.

As a result of this variation in English production, some listeners will lexicalize *tag* with an unreleased plosive, some will lexicalize it with vowel epenthesis. As more people borrow the same word, the two underlying forms will start competing with each other (at the level of surface form), and it is likely that one form wins in the end. Ultimately, the language will end up with some words ending in unreleased plosives, other words ending in epenthesized vowels, and some words may continue to show variation for some time (as both *tag* and *deck* still do, according to Hyunsoon Kim p.c.). Y. Kang (p. 253–4) shows that this gradual elimination of variation indeed happens, and she proposes the mechanism just mentioned (but without mentioning the lexicon).

Apart from the variation in the English plosives, there could also be variation in the rankings of the listener's constraints, as we will see below (25).

Y. Kang (2003) provides explanations for three phonological factors that influence whether vowel insertion appears in Korean adaptation of English words.

First, words with English tense vowels tend to insert a vowel more often than words with English lax vowels: *week* becomes /.wi.kʰɨ./, whereas *quick* becomes /.kʰwik./. Kang (pp. 235–244) argues convincingly that this difference is due to the fact (Parker & Walsh 1981; Y. Kang 2003:239–241) that final consonants in English are more often released after tense than after lax vowels. In our account, this means that the auditory-phonetic input less often contains releases like [ᵏ] after lax vowels than after tense vowels. For instance, the word *quick* was pronounced without a release (i.e. as [_ᵏʰwík˺_]) upon its borrowing into Korean. With Tableau (16) we see that it was perceived as /.kʰwik./. Hence, the form that was adapted is |kʰwik|, which is produced as /.kʰwik./, which sounds as [_ᵏʰːwík˺_] or [_ᵏʰːŷk˺_].

Second, words with English voiced stops in postvocalic word-final position tend to insert a vowel more often than those with voiceless stops: *MIG* becomes /.mi.kɨ./, whereas *kick* becomes /.kʰik./. Kang (pp. 244–249) argues convincingly that this difference is not due to a difference in release frequencies but to the facts that Korean intervocalic lax plosives are phonetically voiced (Kagaya 1974;

Y.Y. Cho 1990; Jun 1995) and that Korean, as does English, lengthens its vowels before voiced consonants (Lim 2000, as reported by Y. Kang 2003:247). The following tableaus, with many of the same constraints as in (10) and (11), show how the reduced voicing cue makes the difference if there is no release burst (similar tableaus can be devised for the lengthening cue):

(24)   Korean perception of the English word *MIG*

[mɪːg˥ ˘˘]	*/+asp ./	*[burst] /C(.)/	*[h] /−asp/	*[no noise] /+asp/	*[˘] /C(.)/	*[˘] /+tense/	*[ ] /ɨ/
/.mik./					*!		
/.mikʰ./	*!			*	*	*	
☞ /.mi.kɨ./							*
/.mi.kʰɨ./				*!		*	*
/.mi.k'ɨ./						*!	*

(25)   Korean perception of the English word *kick*

[_ kʰík˥ _]	*/+asp ./	*[burst] /C(.)/	*[h] /−asp/	*[no noise] /+asp/	*[no voice] /(V)−tense(V)/	*[ ] /ɨ/
☞ /.kʰik./						
/.kʰikʰ./	*!			*		
/.kʰi.kɨ./					*!	*
/.kʰi.kʰɨ./				*!		*
/.kʰi.k'ɨ./						*!

In (24), the cue constraint *[˘]/C(.)/ expresses the idea that if a Korean listener hears even a little bit of voicing, she cannot interpret that as belonging to a final plosive. The existence of Korean forms like /.pik./, borrowed from English *big* (Y. Kang 2003:267), shows that for some borrowers, *[˘]/C(.)/ must be ranked below *[ ]/ɨ/ (this ranking would also be needed to make /.tʰæk./ win in (10)).[15]

Third, words with English dorsal stops in postvocalic word-final position tend to insert a vowel more often than those with labial stops, a fact that Kang attributes to the fact that dorsals are more often released in English than labials are (Rositzke 1943:41; Crystal & House 1988; Byrd 1992). Our model straightforwardly formalizes this explanation with tableaus like (24) and (25). Interestingly, however, coronal plosives are more often borrowed with a release than labial or

---

15.   Candidates like /.mik'./ and /.kʰik'./ are ruled out by a top-ranked */+tense ./ (§2.1).

dorsal plosives, despite the fact that they are less often released than dorsals (Rositzke 1943:41; Crystal & House 1988; Byrd 1992; Y. Kang 2003) or labials (Y. Kang 2003)[16]: *hit* becomes /.hi.tʰɨ./, whereas *tip* and *kick* become /.tʰip./ and /.kʰik./, respectively. Kang explains this special behaviour of coronals as a paradigm uniformity effect related to the alternation that we discuss below in §5; Kang's proposal is plausible, but we do not attempt to give a formalization of this paradigm uniformity effect here.

4.4   Conclusion

In §3 and §4 we have provided an account of all four processes involved in loanword adaptation, without proposing any loanword-specific mechanisms, especially without any loanword-specific ranking of Dep-V. The following section addresses a necessary refinement.

5.   Native alternations in loanwords

In the cases of §4, the underlying forms of the loanwords were completely faithful to the phonological surface forms. This could be expected on the basis of the fact that the only type of constraints involved were faithfulness constraints. In this section we discuss a case where faithfulness is violated, namely the adaptation of English words that end in -*t*. As we saw in §4.3, many of such words are borrowed without vowel insertion, for example /.sjat./ from *shot*. The interesting thing, now, is that these words show signs of ending in an underlying |s|: the accusative of *shot* is /.sja.sɨl./. Thus, the underlying form will be |sjas|, analogously to the native underlying form |os| (§2.1).

In order to be able to handle cases like these, we have to use a more granular set of faithfulness constraints than before. In fact, our set of faithfulness constraints has to express arbitrary relations between underlying and surface form, just as the cue constraints express arbitrary relations between surface form and sound (Fig. 3). First, we make the formulation of Ident(stri) dependent on position, because its ranking may depend on the position. For instance, Ident(stri(.)), which means "in coda position, the underlying and surface values of stridency should be identical", is likely to be ranked lower than its prevocalic counterpart Ident(stri(V)), because stridency faithfulness is especially unimportant in coda position. Next, we

---

16.   Although Kang searched the same database as Byrd (TIMIT), Kang found an opposite difference between coronals and labials than Crystal & House and Byrd. Unlike these other authors, Kang restricted herself to postvocalic plosives, and labelled glottal stops as unreleased coronal plosives.

split up Ident(stri(.)) for its possible arguments /+stri(.)/ and /−stri(.)/, giving the faithfulness constraints *|−stri|/+stri(.)/ and *|+stri|/−stri(.)/. Finally, we include the 'anti-faithfulness' constraints *|+stri|/+stri(.)/ and *|−stri|/−stri(.)/, so that we now have a complete set of arbitrary constraints that link stridency in underlying and surface form.

Of the four constraints, *|+stri|/+stri(.)/ and *|−stri|/+stri(.)/ are of little relevance, given the presence of the high-ranked structural constraint */+stri ./. We are thus left with the two constraints *|−stri|/−stri(.)/ and *|+stri|/−stri(.)/.

The next question is how *|−stri|/−stri(.)/ and *|+stri|/−stri(.)/ are ranked with respect to each other. We observe that underlying final |s|, which always surfaces as /t(.)/, is much more common in Korean than underlying final |t| or |tʰ|. Learning algorithms that are sensitive to frequencies in the data will therefore come to rank *|−stri|/−stri(.)/ over *|+stri|/−stri(.)/ (Boersma 2008). We now show that if we replace the ranking */+stri ./ >> Ident(stri) by the ranking */+stri ./ >> *|−stri|/−stri(.)/ >> *|+stri|/−stri(.)/, we will handle the Korean native production, Korean native recognition, and loanword adaptation correctly.

In production, Tableau (2) turns into Tableau (26), where *|+stri|/−stri(.)/ has taken over the role of Ident(stri) although it is ranked a bit lower.

(26)   L1 Korean production: strident neutralization

|os|	*/+stri ./	Dep-V	Max-C	*|−stri|/−stri(.)/	*/C ./	*|+stri|/−stri(.)/
/.os./	*!				*	
☞ /.ot./					*	*
/.o.sɨ./		*!				
/.o./			*!			

Next, Tableau (27) shows that an underlying final |tʰ| still surfaces as /t./. The requirement is only that *|−stri|/−stri(.)/ is ranked below Max-C.

(27)   L1 Korean production: strident neutralization

|patʰ|	*/+asp ./	*/+stri ./	Dep-V	Max-C	Ident(asp)	*|−stri|/−stri(.)/	*/C ./
/.patʰ./	*!					*	*
☞ /.pat./					*	*	*
/.pa.tʰɨ./			*!			*	
/.pa./				*!			
/.pas./		*!					*

The same constraints are used in recognition. For the native Korean surface form /.pat./, the listener has at least three options for recognition, namely the candidate underlying forms |pas|, |path|, and |patsh|. However, the lexicon links |path| to the morpheme ⟨field⟩, whereas it does not link |pas| or |patsh| to any morpheme. We can express this within a grammar model in which underlying forms are freely generated candidates in an OT tableau (Boersma 2001; Escudero 2005:214–236, Apoussidou 2007:Ch.6). In this model, the relation between underlying forms and morphemes is expressed by *lexical constraints* such as *⟨field⟩|pas| "the morpheme ⟨field⟩ does not link to the underlying form |pas|". As a result, /.pat./ will be recognized as the underlying form |path|, as in Tableau (28).

(28)   Korean recognition of 'field'

/.pat./	*⟨ ⟩	*⟨field⟩ |pas|	*|−stri| /−stri(.)/	*|+stri| /−stri(.)/	*⟨field⟩ |path|
|pas| ⟨ ⟩	*!			*	
☞ |path| ⟨field⟩			*		*
|pas| ⟨field⟩		*!		*	

The first candidate does not violate any lexical constraints, but it links to no morpheme and is therefore ruled out by *⟨ ⟩ (Boersma 2001). The choice between the second and third candidate is handled by the ranking *⟨field⟩|pas| >> *⟨field⟩|path|, which expresses the idea that the Korean morpheme ⟨field⟩ is more strongly connected to the candidate underlying form |path| than to the candidate underlying form |pas|. The tableau assumes that the recognition of the underlying form runs in parallel with the recognition of the morpheme.

All existing words are recognized with the help of lexical constraints, as in (28); for instance, the Korean native sound [ꟸꟸuŋmin] will be unsurprisingly perceived as /.kuŋ.min./, but recognized as the nonfaithful underlying form |kuk+min|; likewise, [ot˺] will be perceived as /.ot./ but recognized as |os|; sometimes, the decision can be made by the lexicon alone, and in other cases (of surface homonymy) syntactic, semantic and pragmatic processing has to be involved;[17] details are outside the scope of the present paper (see the references above). For new loanwords, however, the situation is different: they are not in the lexicon yet, so lexical constraints cannot play a role. As (29) indicates, for example, the only way to recognize /.sjat./ is to link it to no morpheme. The underlying form is then determined by the ranking *|−stri|/−stri(.)/ >> *|+stri|/−stri(.)/.

---

17.   For instance, the surface form /.ot./ could be recognized either as |os| ⟨clothes⟩ or as |otsh| ⟨lacquer⟩. The choice has to be made by higher-level considerations, such as the pragmatic context, which are not modelled here.

(29)   Korean recognition of 'shot'

/.sjat./	*⟨ ⟩	*⟨clothes⟩ |otʰ|	*|−stri| /−stri(.)/	*|+stri| /−stri(.)/	*⟨clothes⟩ |os|
|sjatʰ| ⟨ ⟩	*		*!		
☞ |sjas| ⟨ ⟩	*			*	

The winning candidate |sjas| is thus ultimately determined by frequency: the frequency-dependent ranking of faithfulness constraints causes loanword adapters to posit the underlying final segment that most frequently corresponds to it in the rest of the vocabulary.[18] The learner can subsequently create a new lexical item |sjas| ⟨shot⟩. In (29), we finally see the reason for splitting up the faithfulness constraints.

6.   Comparison to other models

Our model explains both the auditory similarity and the differences between the forms of the donor language and the borrowing language. Auditory similarity is achieved by the bidirectionality of the cue constraints (§4.2) and to a lesser extent the bidirectionality of the faithfulness constraints (§5); differences occur as a result of crosslinguistic differences in the rankings of cue constraints, which affect loanword perception (§3.1–2) as well as loanword production (§4.2), and differences in the rankings of structural constraints, which affect loanword perception (§3.3). In this section we discuss how other authors have handled Korean loanwords within their models, or how they probably would have handled them if they had discussed Korean within their models. It turns out that by regarding perception as a less active process than we do, all these models have had to posit and incorporate loanword-specific devices.

6.1   The "all phonology is production" assumption

Many authors assume that loanword adapters store the donor language's phonetic or surface form more or less directly as an underlying form in the receiving language, and that subsequently, the (production) grammar performs the adaptation to the native phonology. The role of perception in the first step (the storage

---

18.   This does not happen only to loanword adapters. Albright (2002:112) mentions that Korean is going through a change in which the most frequent underlying forms that correspond to surface final /t/, namely |s| and |tsʰ|, are taking over *native* paradigms with original underlying |tʰ| (and |ts|), sometimes piecemeal. Thus, next to the locative /.pa.tʰɛ./ we find topic forms such as /.pa.sɨn./ and /.pa.tsʰɨn./.

process) is either absent (Paradis & LaCharité 1997), or restricted to a limited number of extragrammatical adaptations to the segmental or tonal inventories of the receiving language (Silverman 1992; Yip 1993). The role of perception in the second (production) step is either absent, or reflected in the ranking of faithfulness constraints (Steriade 2001).

In these views, perception is therefore extragrammatical and only indirectly influences the production. Maintaining such a view turns out to run into several problems, such as loanword-specific constraints, loanword-specific rankings, or failures to handle the data. In the following paragraphs we discuss several specific approaches.

For the Korean case, H. Kang (1996) assumes that the English word *stress* is stored as the underlying form |stʰrɛs'| (in our notation). After this, the phonology converts this |stʰrɛs'| to the surface form /.sɨ.tʰɨ.rɛ.s'ɨ./. Kang therefore concludes that in loanword phonology, IDENT(stri) outranks DEP-V (otherwise the surface form would have ended in /t./). Meanwhile, Kang notices that an underlying native |os| surfaces as /.ot./. Kang therefore concludes that in the native phonology, the anti-vowel-insertion constraint DEP-V outranks the faithfulness constraint IDENT(stri). In other words, the same constraints are ranked differently in loanword adaptation than in native phonology (this is a typical problem in the loanword literature: also Itô & Mester 1995; Shinohara 2004). In our model, no such loanword-specific rankings are required: vowel insertion is allowed in perception by the ranking *[*friction*]/−stri/ >> *[ ]/ɨ/ (§3.3), following the general observation that listeners routinely have to work with missing cues (§3.2), whereas vowel insertion is disallowed in production by the ranking DEP-V >> IDENT(stri) (§2.1).

Production-based accounts often involve storing phonetic detail in underlying forms; a high ranking of faithfulness then forces this detail to the surface. For Korean, Y. Kang (2003:253) states that the English word *jeep* is borrowed with a phonetically detailed underlying form, namely variably (in our notation) as |_ᵗˢijp⌐_ᵖʰ| (with a release) or as |_ᵗˢijp⌐_| (without a release). A single constraint ranking then maps e.g. |_ᵗˢijp⌐_ᵖʰ| to the surface form /.tsi.pʰɨ./ (or [_ᵈᶻˌɪp⌐_ᵖʰˑɨ]; Kang makes no difference between phonetic and surface form). As a result, Kang states that a faithfulness constraint like MAX[*release*] has to outrank DEP-V. There are two problems with this proposal. First, it is usual in phonology to regard underlying forms as economical representations without phonetic detail. Second, it contains a contradiction: although Kang explicitly states that vowel insertion takes place in perception (as we acknowledged throughout §3 and §4), her proposed underlying forms do not contain any inserted vowels. This is a contradiction because psycholinguistic models of speech comprehension generally state that lexical representations have been filtered by the perception process (Cutler & McQueen 1997; Peperkamp & Dupoux 2003, also our Fig. 1), so that if the perception process inserts

a vowel, this vowel should end up in the underlying form as well. Kang's OT proposal can therefore be repaired by formalizing the observed loanword adaptation in two steps: first the perception of the sound [_^{dʒ}ɪ⁻jp⁻^{ph}] as the surface structure /.tsi.pʰɨ./ (formalized with OT perception tableaus), followed by the production of the resulting underlying form |tsipʰɨ| as /.tsi.pʰɨ./ and [_^{dʑ}ɪp⁻_^{pʰi}ɨ] (formalized with OT production tableaus). This is what we have done in §3 and §4; no phonetic detail appears in the economical underlying form, and the underlying form honours all filterings by the perception process, including the insertion of a vowel.

The existence of a high-ranked MAX (often with a low-ranked DEP) in loanword adaptation has been raised to the status of a "preservation principle", according to which elements that are present in the form provided by the donor language tend to survive in the receiving language (Paradis & LaCharité 1997, 2005; Rose & Demuth 2006). In our model, this "principle" has a direct explanation in terms of high-ranked cue constraints for positive auditory cues (and concomitant low-ranked cue constraints against inserting 'illusory' phonological material) (Boersma 2007a: Footnotes 26 and 27).

Some production-based accounts reject loanword-specific rankings for the same reasons as we do. However, such accounts often require loanword-specific *constraints* that ensure faithfulness to the auditory information of the donor language. Examples are Davidson & Noyer's (1997) constraint MATCH, Kenstowicz's (2005) phonetic output-output faithfulness, and Yip's (2006) constraint MIMIC. To account for the Korean facts, these models would indeed require such constraints (as exemplified by Kenstowicz 2005:§3.1; also Smith 2006 for Japanese), because these models still regard perception as at most a passive low-level extragrammatical device that allows only the interpretation of nonnative sounds in terms of the native phoneme inventory, and deletions in case of poor audibility. These models cannot handle perceptual insertion, because that would require an integration of perception and phonology, as we have shown.

Within the tradition started by Silverman, only Peperkamp & Dupoux (2003; followed by Iverson & Lee 2006) agree that perception can insert vowels. They propose that all loanword adaptations take place in an extragrammatical perception module, and that the set of adaptations includes not just Silverman's mapping to native segments and tones, but also a mapping to native syllables, which allows insertion. Peperkamp & Dupoux' proposal cannot handle, though, the difference between /.tsʰæp.tʰʌ./ and /.pʰi.kʰɨ.nik./, which is due to a phonotactic (*phonological*) constraint and cannot be regarded as a difference in syllable perception.

To sum up: all these authors assume that phonologically-informed adaptations can only be made in production, and that perception is a passive process (also Hall & Hamann 2003; Miao 2005; Kenstowicz & Suchato 2006; Davis & Cho 2006; Adler 2006; Uffmann 2006; explicitly: Shinohara 2006). However, to account

for the Korean facts without loanword-specific rankings or constraints, one has to acknowledge instead that phonologically-informed loanword adaptation occurs to a large part in perception (for exceptions see §7.3), and that therefore perception is just as phonological as production, as it is in the L1-based model of Fig. 1.

6.2   Perception is phonological as well

Perception, then, is phonological itself. That is, vowels are inserted in Koreans' perception of English words partly because alternative candidates violate Korean-specific phonological constraints.

H. Kim (this volume) provides a non-OT account of Korean loanwords in which she uses a 'feature-driven' model: a first 'perception' stage (following Peperkamp & Dupoux) matches the English auditory cues to Korean-specific features and syllables and can therefore insert vowels; in a second 'grammar' stage, still in the comprehension direction, structural constraints can exert their influence. Of the latter, Kim gives an example of a phonotactic restriction against homorganic glide-vowel sequences like */jɛ/, which causes English /ʃ/ to be borrowed before front vowels as /sw/ (/.swɛl./ 'Shell') instead of as the auditorily preferred /sj/. We agree with both types of influence that Kim proposes. As we have seen, though, in the examples of English complex onsets and of /.tsʰæp.tʰʌ./ versus /.pʰi.kʰɨ.nik./, vowel insertion (in Kim's first stage) is itself influenced by phonotactic restrictions (Kim's second stage), so it seems to be more parsimonious to model them in a single perception stage, as can be done in Optimality Theory, as we have shown.[19]

That the same phonotactic restrictions influence perception as well as production does not mean that the *repair* of the forbidden phonological structure is the same in perception and production. Thus, the forbidden structure /k.m/ is repaired as /kɨ.m/ in perception (§3.3) but as /ŋ.m/ in production (§2.3). This asymmetry was noted by Kabak & Idsardi (2007), who concluded that "native phonological rules" (nasalization of /k/ in production) do not "affect the perceptual processing" of strings like [km] (p. 48). It is indeed not the native phonological *rule* (nasalization of /k/) that affects the perceptual processing of [km], but the native phonological *constraint* (the syllable contact law) that affects the perceptual processing of [km], namely by inserting a vowel. The possibility of having the same constraint but different kinds of repairs is typical of analyses in Optimality Theory, and it is

---

19.   Another concern with Kim's model is that the grammar has to influence perception in the first stage as well, since it restricts the inventory of phonological elements that build the output of perception (as Kim states explicitly). Having the language-specific handling of cues interact with language-specific phonotactic constraints in parallel, as is done in the present paper, automatically alleviates this concern.

therefore our use of OT in modelling perception as well as production that led us to regard the perception of [km] as /kɨ.m/ and the production of |km| as /ŋ.m/ as two diffferent outcomes of the same phonological restriction.

We are not the only ones who have attempted to model both perception and production in OT. Kenstowicz (2001) proposes that in loanword adaptation, the "loan source" is first filtered through an OT "perception grammar", which results in a "lexical representation". This lexical representation (underlying form) is then filtered by an OT "production grammar", which results in the "output". However, as Silverman and Yip, Kenstowicz regards vowel insertion as a task of the production grammar; the nature of the inserted material is then determined by the principle of "minimal saliency" (Shinohara 1997:Fn.32, Steriade 2001:238; also Kenstowicz 2003). A more general problem is that although Kenstowicz uses the term "perception grammar", he regards this as a direct mapping from sound to underlying form, unlike Boersma (1998), who introduced this term as the first step in a two-stage comprehension process (here, Fig. 1). This means that Kenstowicz would often have to propose (as he does) that constraint rankings are different in comprehension than in production.

Some authors agree with us (and with Polivanov 1931) that the lexicon can contain underlying forms that have been filtered by the perception process on the basis of language-specific structural restrictions (and not just segmental similarity). Broselow (2004, to appear) does propose structural constraints in comprehension: Broselow (2004) proposes that the "perception grammar" contains the strong constraint "any stressed foot is followed by a word edge", which causes the Spanish form [garabáto] to be stored in the Huave lexicon as |garabát|. In a new version of her paper, Broselow (to appear) extends this view to vowel insertion (without formalization): she reports with approval a proposal by Schütz (1978) that Fiji listeners of English indeed hear the word *whiskey* as /.wi.si.ki./ (as Y. Kang 2003 does for the Korean case, Schütz reportedly relies on arguments of releases and vowel degradation).

However, just as Kenstowicz, Broselow regards the perception grammar (contra Boersma 1998) as being a direct mapping from sound to underlying form. Such a two-level view of representations poses a general problem. In Broselow's proposal, structural constraints can only apply to the output of the entire comprehension process, i.e. to the underlying form, and this is indeed what she proposes. But it is usual in phonology to regard underlying forms as economical representations that are devoid of structures such as feet, syllables, and codas. Even for Broselow's own constraint "any stressed foot is followed by a word edge", this is already problematic, because the underlying form does not contain any feet; instead, feet are properties of phonological surface structures, so the comprehension mapping should be [garabáto] → /gara(bát)/ → |garabát|, and the relevant

constraint should be the cue constraint "an auditorily stressed vowel should be perceived as being final in its foot". For the general case, a two-level account such as Broselow's would only be possible if all structural constraints could refer to underlying material (segments, word boundaries) alone, and not to metrical elements. For Korean, the constraints would have to refer to structures like #CC, CCC, and CC#, thereby losing the generalization that Korean phonotactics can be expressed in terms of the simple syllable structure constraints */.CC/ and */CC./. A remarkable feature of Broselow's proposal is that the output of production does have metrical structure, whereas the input to comprehension has not (it is more 'phonetic'); hence, these two 'surface' representations do not seem to be the same; in some sense, then, Broselow's model does seem to require three different representations; it seems to be only a small step to conclude that all three representations must play a role in both comprehension and production, as they do in the older model of Fig. 1.

We conclude that a full account of loanword adaptation requires the same three levels of representation that a full account of L1 phonology and phonetics requires. A three-level model allows us to work with structural constraints on surface forms, both in production and in perception.

7.   Discussion

In this section we discuss a number of remaining issues, and end by giving a complete ranking.

7.1   Phonology without structural constraints

One could argue that structural constraints are not necessary for formalizing perception, because their effects can equally well be described with the right number of cue constraints. This is true. The structural constraint */+stri ./ could be replaced with a large number of cue constraints such as *[s]/+stri ./, *[tʰ]/+stri ./ and *[bzɪltʃ]/+stri ./; and if we simplifyingly write SYLLCON as the structural constraint */k.n/, we see that we can replace it with a large number of cue constraints *[X]/k.n/, where [X] is any auditory form. Replacing structural with cue constraints in this way gives up on the formalization of generalizations, which is why we did not do it in this paper; but we cannot deny that it is *possible*.

From these observations, one might argue that perception and production require different formalizations, because the formalization of production does require structural constraints, and the formalization of perception could be (awkwardly)

performed with a massive number of cue constraints. However, the fact is that once all these cue constraints exist, they can handle production as well! This is illustrated in Tableau (30).

(30)   L1 Korean production without structural constraints

|os|	*[s] /+stri ./	*[t˺] /+stri ./	Dᴇᴘ-V	Mᴀx-C	Iᴅᴇɴᴛ(stri)
/.os./[os]	*!				
/.os./[ot˺_]		*!			
☞ /.ot./[ot˺_]					*
/.o.sɨ./[osɨ]			*!		
/.o./[o]				*!	

In (30), phonological production (the mapping from underlying to surface form) and phonetic implementation (the mapping from surface to phonetic form) are handled in parallel, as in Boersma (2007ab, 2008); that is, the candidates are pairs of surface and phonetic form, and faithfulness and cue constraints interact with each other. We see that the surface structure /.os./ is ruled out by the cue constraints *[s]/+stri ./ and *[t˺]/+stri ./, and that no structural constraints are needed.

We conclude that if structural constraints can be replaced with cue constraints for formalizing perception, they can also be replaced with cue constraints for formalizing production (at least if the phonology and the phonetics are handled in parallel). Therefore: whether or not we use structural constraints for describing linguistic processes, perception and production will be modelled as equally 'phonological'.

7.2   Nonnative phonotactics in loanword adaptation

It has been observed that loanwords often introduce nonnative phonotactics (Haugen 1950:217,226; for Korean liquids: Kenstowicz 2005:24). For instance, the English word *shot* is borrowed into Korean as the surface form /.sjat./ despite the fact that in the native Korean vocabulary syllables rarely start with /.sj/ and the sequence /ja/ is rarely preceded by an onset consonant. Apparently, Korean has the structural constraints */.sj/ and */.Cja/. How then is it possible that *shot* is borrowed as /.sjat./?

The answer is that in perception it is possible (given factorial typology) that some cue constraints outrank some structural constraints. Tableau (31) shows the interaction for *shot*.

(31)   Korean perception of the English word *shot*

[ʃɒtˀ_]	*[friction] / /	*[high F2] / /	Max-C	*/.sj/	*/.Cja/
/.sat./		*!			
/.jat./	*!				
☞ /.sjat./				*	*

Perceiving [ʃɒtˀ_] as /.sat./ would ignore the auditory event [*high F2*], which in Korean is a strong cue in favour of the palatal segments /j/ or /i/. Perceiving [ʃɒtˀ_] as /.jat./ would ignore the auditory event [*friction*], which in Korean is a strong cue in favour of the sibilant consonants /s/ or /ts/. Thus, the perceived form is /.sjat./, the underlying form, mentioned in (29), is |sjas|, and the produced form is /.sjat./, assuming that Max-C outranks the two structural constraints in (31).

We conclude that regarding perception as an interaction between structural and cue constraints predicts the existence of nonnative phonotactics in loanword adaptation, which is indeed attested. According to computer simulations of a bidirectional learning algorithm (Boersma 2008), cue constraints are expected to be ranked high if confusability is low (i.e. if auditory salience is high or the native language lacks confusing phonological competitors) or if the phonological element's frequency in the native language is moderately low.

7.3   Loanword adaptation that takes place outside perception

In this paper we have focussed on cases of loanword adaptation that take place in perception. The model of Fig. 1 predicts that there are several other processes in which loanword adaptation can take place.

One of those processes is recognition, i.e. the mapping from phonological surface form to underlying form. We saw an example of adaptation in recognition in (29), where a final /t/ in the surface form was installed in the lexicon as a final |s|. The phonological production process, i.e. the mapping from underlying to surface form, then causes this |s| to appear as /s/ in the accusative /.sja.sɨl./.

Another process is phonetic implementation, i.e. the mapping from phonological surface form to phonetic form. We saw an example of adaptation in phonetic implementation in (21) and (22), where the Korean ranking of cue constraints ensured that English loanwords are pronounced with a Korean accent.

Beside perception, recognition, and production, there may be other sources of loanword adaptation. Orthography has been claimed to have introduced the form /.pʰiŋ.nik./ into Korean (Kabak 2003:59). In Fig. 1 this would be viewed as an interpretation of the English spelling *picnic* in terms of the two Korean syllabic characters 픽 닉, which are then mapped to the Korean underlying

form |pʰik-nik|, which is subsequently produced as /.pʰiŋ.nik./ by the rules of Korean phonology.

7.4   Loanword adaptation by bilinguals

It is likely that loanword adaptation is partly performed by advanced L2 speakers (Paul 1880; Haugen 1950; Paradis & LaCharité 1997; LaCharité & Paradis 2005). If this occurs, English loanwords may be filtered by L2 English perception rather than by native Korean perception, because L2 listeners have been found to shift their perceptual boundaries depending on the language they think they hear (Elman, Diehl & Buchwald 1977; Boersma & Escudero 2008). Also, lexical storage may occur in terms of an L2 English inventory rather than in terms of the native Korean inventory, because L2 listeners have been found to reuse their native inventories in lexical representations (Boersma & Escudero 2008). For example, bilinguals may analyse English as having only lax plosives such as /p/ and aspirated plosives such as /pʰ/, and therefore as lacking /p'/. This means that the English word *bye* is interpreted as starting with /p/ and *pie* as starting with /pʰ/. For the labial plosive in *spy* they would have two options; if the voicing cue weighs heavier than the aspiration cue, they will interpret the plosive as /pʰ/. This may be the explanation behind the aspirated plosive that appears in loanwords like |sɨpʰaikʰɨ| (as well as the avoidance of /.pʰi.k'ɨ.nik./ mentioned in Footnote 12). In subsequent production in Korean, the bilinguals will then use the native Korean grammar. In an L2 version of the model of Fig. 1, comparable facts have been accounted for by modelling the acquisition of L2 underlying forms with a morpheme-driven learning algorithm (Escudero 2005:214–236; Weiand 2007); for English *spy*, there would be a long-lasting competition between |sɨpʰai| and |sɨp'ai|, which would ultimately be won by |sɨpʰai| because of its more peripheral auditory correlates (this confirms a hypothesis by Kenstowicz 2005, although it does not require his formalization in terms of MɪɴDɪsᴛ constraints). As a result of the need to map [sp-] to |sɨpʰ-|, Korean bilinguals will adapt their perception of English in such a way that the boundary between their L2 intervocalic /p/ and /pʰ/ falls in between that of the English *bye* and *spy* (thus, the cue constraints for the voice-onset-time continuum will be ranked differently in L1 and L2). The same mechanism could help L2 learners to equate the English /æ/-/ɛ/ contrast with the Korean /æ/-/ɛ/ contrast, despite the acoustic differences (§3.1). Modelling these facts would require computer simulations such as those performed by Weiand.

7.5   The complete grammar

The grammar in (32) combines all the constraints that we used in this paper, except the 20 cue constraints for the front vowels in §3.1, and the specific lexical constraints of §5. The ranking is a possible division into five strata.

(32) Native Korean grammar, which also accounts for the adaptation of English words

{　　*/+asp ./　Dᴇᴘ-V　SʏʟʟCᴏɴ　*[*burst*]/C(.)/　*[◌̰]/+asp/　*[ʰː]/−asp/
　　　　*/+tense ./　*/+stri ./　*/CC ./　　*/.CC/　　*[C˺]/ /
　　　　*[*friction*]/−stri/　　*[*friction*]/ /　　*[_]/+nas/　　*⟨ ⟩　*[*high F2*]/ / }
　　　　　　　　　>>

{　　Mᴀx-C　*[*no noise*]/+asp/　　*[ʰ]/−asp/　　*[ˉ]/+tense/　　*[ʹ]/−tense/
　　　　*[◌̰]/C(.)/　*[◌̰]/+tense/　　*[*no voice*]/(V)−tense(V)/　　}
　　　　　　　　　>>

{　　Iᴅᴇɴᴛ(asp)　Iᴅᴇɴᴛ(stri(V))　　*|−stri|/−stri(.)/　　Iᴅᴇɴᴛ(nas)
　　　　*[◌̆]/C(.)/　　*[◌̆]/+tense/　　*/C ./　　*/.sj/　　*/.Cja/　　}
　　　　　　　　　>>

{　　*[ ]/ɨ/　　*[ʰ]/+asp/　　(also e.g. *[*no noise*]/−asp/)
　　　　*|+stri|/−stri(.)/　　articulatory constraints　}
　　　　　　　　　>>

{　　*[ʰː]/+asp/　　(also e.g. *[◌̰]/−asp/)　　}

All constraints in this single ranking (except those in the bottom stratum, which could be removed from the grammar) are needed for comprehending and/or producing Korean in everyday use. The very same ranking also explains all the loanword adaptation phenomena that we discussed in this paper.

We like to stress that the ranking in (32), despite its size, does not contain any unlikely rankings or interaction tricks: the structural and faithfulness constraints are ranked on the basis of relatively uncontroversial facts of Korean phonological production (although the ranking works for comprehension as well), and the cue constraints are ranked as follows: constraints against strengthened 'contrary' cues at the top, constraints against 'normal' contrary cues in the second stratum, constraints against 'reduced' contrary cues in the third stratum, constraints against 'friendly' cues in the fourth stratum, and constraints against 'very friendly' cues in the bottom stratum. We also note that the set of cue constraints in (32), despite its size, is still rather minimalistic: a full set would require much finer-grained auditory distinctions. For instance, we ignored in this paper the fact that the lax plosives are slightly aspirated, that the aspirated plosives tend to have a higher F0 than the tense plosives, that aspirated plosives are less aspirated in intervocalic position, and that lax plosives are shorter than tense plosives.

8. Conclusion

In the present paper we have applied an existing bidirectional model of L1 phonology and phonetics (Fig. 1) to several cases of loanword adaptation in Korean. By regarding perception as equally phonological as production, this L1 model turns

out to handle the loanword adaptation facts without assuming any additional (i.e. loanword-specific) rankings, constraints, or other devices. Instead, loanword adaptation is fully explained by the behaviour of listeners in their native language. As a side effect, we have reconciled the phonology-versus-perception debate in loanword adaptation research: perception simply *is* phonological. The assumptions that have proven crucial for achieving this result (all visible in Fig. 1) are the distinction between phonological and phonetic representations, the bidirectionality of cue and faithfulness constraints, and the use of structural constraints both in perception and production. All these assumptions have proven necessary for L1 phonology as well (Boersma 2007ab, 2008; Boersma & Hamann 2008) and are therefore not specific to loanword adaptation.

By doing away with loanword-specific phonology, we hope to have reduced the mystery of loanword adaptation. Korean has provided many interesting examples, and our model handled all of them in a straightforward way. It will be interesting to see how our model performs on languages that might exhibit types of loanword adaptations that we did not discuss.

References

Adler, Allison N. (2006). Faithfulness and perception in loanword adaptation: a case study from Hawaiian. *Lingua* 116:1024–1045.

Ahn, Sang-Cheol & Gregory K. Iverson (2004). Dimensions in Korean laryngeal phonology. *Journal of East Asian Linguistics* 13:345–379.

Albright, Adam (2002). *The identification of bases in morphological paradigms*. Ph.D. dissertation, UCLA.

Apoussidou, Diana (2007). *The learnability of metrical phonology*. Ph.D. dissertation, University of Amsterdam.

Boersma, Paul (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21:43–58.

Boersma, Paul (1998). *Functional phonology: formalizing the interaction between articulatory and perceptual drives*. Ph.D. dissertation, University of Amsterdam.

Boersma, Paul (2000). The OCP in the perception grammar. *Rutgers Optimality Archive* 435.

Boersma, Paul (2001). Phonology-semantics interaction in OT, and its acquisition. In: Robert Kirchner, Wolf Wikeley & Joe Pater (eds.) *Papers in Experimental and Theoretical Linguistics, Vol. 6*. Edmonton: University of Alberta, 24–35.

Boersma, Paul (2006). Prototypicality judgments as inverted perception. In: Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.) *Gradience in grammar*. Oxford: Oxford University Press, 167–184.

Boersma, Paul (2007a). Some listener-oriented accounts of *h-aspiré* in French. *Lingua* 117:1989–2054.

Boersma, Paul (2007b). Cue constraints and their interactions in phonological perception and production. *Rutgers Optimality Archive* 944.

Boersma, Paul (2008). Emergent ranking of faithfulness explains markedness and licensing by cue. *Rutgers Optimality Archive* 954.

Boersma, Paul & Paola Escudero (2008). Learning to perceive a smaller L2 vowel inventory: an Optimality Theory account. In: Peter Avery, B. Elan Dresher & Keren Rice (eds.) *Contrast in phonology: theory, perception, acquisition*. Berlin: Mouton de Gruyter, 271–301.

Boersma, Paul, Paola Escudero & Rachel Hayes (2003). Learning abstract phonological from auditory phonetic categories: an integrated model for the acquisition of language-specific sound categories. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona: 1013–1016.

Boersma, Paul & Silke Hamann (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25:217–270.

Broselow, Ellen (2004). Language contact phonology: richness of the stimulus, poverty of the base. *Proceedings of the North-Eastern Linguistic Society* 34:1–22.

Broselow, Ellen (to appear). Stress adaptation in loanword phonology: perception and learnability. In: Paul Boersma & Silke Hamann (eds.) *Phonology in perception*. Berlin: Mouton de Gruyter.

Byrd, Dani (1992). A note on English sentence-final stops. *UCLA Working Papers in Phonetics* 81:37–38.

Cho, Taehong, Sun-Ah Jun & Peter Ladefoged (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30:193–228.

Cho, Young-mee Yu (1990). Syntax and phrasing in Korean. In: Sharon Inkelas & Draga Zec (eds.) *The phonology-syntax connection*. Chicago: University of Chicago Press, 47–62.

Cruttenden, Alan (1994). *Gimson's pronunciation of English*. London: Arnold.

Crystal, Thomas H. & Arthur S. House (1988). Segmental durations in connected-speech signals: current results. *Journal of the Acoustical Society of America* 83:1553–1573.

Davidson, Lisa & Rolf Noyer (1997). Loan phonology in Huave: nativization and the ranking of faithfulness constraints. *West Coast Conference on Formal Linguistics* 15:65–79.

Davis, Stuart & Mi-Hui Cho (2006). Phonetics versus phonology: English word final /s/ in Korean loanword phonology. *Lingua* 116:1008–1023.

Davis, Stuart & Seung-Hoon Shin (1999). The syllable contact constraint in Korean: an Optimality-theoretic analysis. *Journal of East Asian Linguistics* 8:285–312.

Dupoux, Emmanuel, Kazuhiko Kaheki, Yuki Hirose, Christophe Pallier & Jacques Mehler (1999). Epenthetic vowels in Japanese: a perceptual illusion. *Journal of Experimental Psychology: Human Perception and Performance* 25:1568–1578.

Elman, Jeffrey l., Randy L. Diehl & Susan E. Buchwald (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America* 62:971–974.

Escudero, Paola (2005). *Linguistic perception and second language acquistion: explaining the attainment of optimal phonological categorization*. Ph.D. dissertation, Utrecht University.

Escudero, Paola & Paul Boersma (2003). Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. In: Sundha Arunachalam, Elsi Kaiser & Alexander Williams (eds.) *Proceedings of the 25th Annual Penn Linguistics Colloquium (Penn Working Papers in Linguistics 8.1)*, 71–85.

Escudero, Paola & Paul Boersma (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26:551–585.

Flemming, Edward (1995). *Auditory representations in phonology*. Ph.D. dissertation, UCLA.

Hall, T.A. & Silke Hamann (2003). Loanword nativization in Modern Standard German as evidence for the phonological word. *Zeitschrift für Sprachwissenschaft* 22:56–85.

Hamann, Silke (2009). Variation in the perception of an L2 contrast: a combined phonetic and phonological account. In: Frank Kügler, Caroline Féry & Ruben van der Vijver (eds.) *Variation and gradience in phonetics and phonology*. Berlin: Mouton de Gruyter, 79–105.

Han, Mieko S. & R.S. Weitzman (1970). Acoustic features of Korean /P, T, K/, /p, t, k/ and /ph, th, kh/. *Phonetica* 22:112–128.

Hardcastle, William J. (1973). Some observations of the tense-lax distinction in initial stops in Korean. *Journal of Phonetics* 1:263–271.

Haugen, Einar (1950). The analysis of linguistic borrowing. *Language* 26:210–231.

Heffner, Roe-Merrill Secrist (1937). Notes on the length of vowels. *American Speech* 12:128–134.

Hirose, Hajime, Chao-Yang Lee & Tatsujiro Ushijima (1974). Laryngeal control in Korean stop production. *Journal of Phonetics* 2:145–152.

Hooper, Joan B. (1976). *An introduction to Natural Generative Phonology*. New York: Academic Press.

House, Arthur S. & Gordon H. Fairbanks (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25:105–113.

Hwang, Jiwon, Ellen Broselow, Susana de Leon, Nacy Squires & Crystal Sandiford (2007). Asymmetries in perception and production. Paper presented at the Experimental Optimality Theory Workshop, Ann Arbor, 19 May 2007.

Ingram, John C. & See-Gyoon Park (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics* 25:343–370.

Ito, Chiyuki, Yoonjung Kang & Michael Kenstowicz (2006). The adaptation of Japanese loanwords into Korean. *MIT Working Papers in Linguistics* 52:65–104.

Itô, Junko & Armin Mester (1995). The core-periphery structure of the lexicon and constraints on reranking. In: Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyck (eds.) *Papers in Optimality Theory (University of Massachusetts Occasional Papers 18)*. Amherst, Mass.: Graduate Linguistics Student Association, 181–209.

Iverson, Gregory K. (1983). On glottal width features. *Lingua* 60:331–339.

Iverson, Gregory K. & Ahrong Lee (2006). Perception of contrast in Korean loanword adaptation. *Korean Linguistics* 13:49–87.

Iverson, Gregory K. & Hyang-Sook Sohn (1994). Liquid representation in Korean. In: Young-Key Kim-Renaud (eds.) *Theoretical issues in Korean linguistics*. Stanford: Stanford University Centre for the Study of Language and Information, 79–100.

Jun, Sun-Ah (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In: Bruce Connell & Amalia Arvaniti (eds.) *Phonology and phonetic evidence: Papers in laboratory phonology*. Cambridge: Cambridge University Press, 235–253.

Kabak, Baris (2003). *The perceptual processing of second language consonant clusters*. Ph.D. dissertation, University of Delaware.

Kabak, Baris & William Idsardi (2007). Perceptual distortions in the adaptation of English consonant clusters: syllable structure or consonantal contact constraints? *Language and Speech* 50:23–52.

Kagaya, Ryohei (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics* 2:161–180.

Kang, Hyunsook (1996). English loanwords in Korean. *Studies in Phonetics, Phonology and Morphology* 2:21–47.

Kang, Hyunsook (1999). Affixation to English loanwords. *Studies in Phonetics, Phonology and Morphology* 5:291–307.

Kang, Hyunsook (2002). On the Optimality-theoretic analysis of Korean nasal-liquid alternations. *Journal of East Asian Linguistics* 11:43–66.

Kang, Yoonjung (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20:219–273.

Kenstowicz, Michael (2001). The role of perception in loanword phonology. *Linguistique Africaine* 20.

Kenstowicz, Michael (2003). Salience and similarity in loanword adaption: a case study from Fijian. *Rutgers Optimality Archive* 609.

Kenstowicz, Michael (2005). The phonetics and phonology of Korean loanword adaptation. In: Sang-Jik Rhee (eds.) *Proceedings of the first European conference on Korean linguistics*. Seoul: Hankook Publishing Co., 17–32.

Kenstowicz, Michael & Atiwong Suchato (2006). Issues in loanword adaptation: a case study from Thai. *Lingua* 116:921–949.

Kim, Hyunsoon (1998). A phonetic characterization of release and nonrelease: the case of Korean and English. *Language Research* 34:347–368.

Kim, Hyunsoon (2005). The representation of the three-way laryngeal contrast in Korean consonants. In: Marc van Oostendorp & Jeroen van de Weijer (eds.) *The internal organization of phonological segments*. Berlin: Mouton de Gruyter, 263–293.

Kim, Hyunsoon (this volume). Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation. In: Leo Wetzels & Andrea Calabrese (eds.) *Studies in Loan Phonology*. Amsterdam: Benjamins.

Kim, Mi-Ran Cho (1994). *Acoustic characteristics of Korean stops and perception of English stop consonants*. Ph.D. dissertation, University of Wisconsin, Madison.

Kim, Mi-Ryoung (2000). *Segmental and tonal interactions in English and Korean: a phonetic and phonological study*. Ph.D. dissertation, University of Michigan.

Kim, Mi-Ryoung, Patrice Speeter Beddor & Julie Horrocks (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30:77–100.

Kim, Mi-Ryoung & San Duanmu (2004). "Tense" and "lax" stops in Korean. *Journal of East Asian Linguistics* 13:59–104.

Kim-Renaud, Young-Key (1977). Syllable-boundary phenomena in Korean. *Korean Studies* 1:243–273.

Kim-Renaud, Young-Key (1987). Fast speech, casual speech and restructuring. *Harvard Studies in Korean Linguistics* 2:341–359.

Kirchner, Robert (1998). *An effort-based approach to consonant lenition*. Ph.D. dissertation, UCLA.

LaCharité, Darlene & Carole Paradis (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36:223–258.

Lee, Iksop & Robert Ramsey (2000). *The Korean language*. New York: State University of New York Press.

Lehiste, Ilse & Gordon E. Peterson (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America* 33:419–425.

Levelt, Willem J.M. (1989). *Speaking: from intention to articulation*. Cambridge, Mass.: MIT Press.

Lim, Byung-jin (2000). Effects of syllable structure and syllable boundary on segmental duration in Korean. *Journal of the Acoustical Society of America*, 107:2804.

Lisker, Leigh & Arthur S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20:384–422.

Martin, Samuel E. (1951). Korean phonemics. *Language* 27:519–533.

McCarthy, John J. & Alan Prince (1995). Faithfulness and reduplicative identity. In: Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyck (eds.) *Papers in Optimality Theory (University*

of Massachusetts Occasional Papers 18). Amherst, Mass.: Graduate Linguistics Student Association, 249–384.

McQueen, James & Anne Cutler (1997). Cognitive processes in speech perception. In: William J. Hardcastle & John Laver (eds.) *The handbook of phonetic sciences*. Oxford: Blackwell, 566–586.

Miao, Ruiqin (2005). *Loanword adaptation in Mandarin Chinese: perceptual, phonological and sociolinguistic factors*. Ph.D. dissertation, Stony Brook University.

Murray, R. & Theo Vennemann (1983). Sound change and syllable structure: problems in Germanic phonology. *Language* 59:514–528.

Oh, Mira (1996). Linguistic input to loanword phonology. *Studies in Phonetics, Phonology and Morphology* 2:117–126.

Ohde, Ralph N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America* 75:224–230.

Paradis, Carole & Darlene LaCharité (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33:379–430.

Paradis, Carole & Darlene LaCharité (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36:223–258.

Park, Hanyong (2007). Varied adaptation patterns of English stops and fricatives in Korean loanwords: the influence of the P-map. *Indiana University Linguistics Club Working Papers* 7.

Parker, Frank & Thomas Walsh (1981). Voicing cues as a function of the tense/lax distinction of vowels. *Journal of Phonetics* 9:353–358.

Pater, Joe (2004). Bridging the gap between perception and production with minimally violable constraints. In: René Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in phonological acquisition*. Cambridge: Cambridge University Press, 219–244.

Paul, Hermann (1880). *Principien der Sprachgeschichte*. Halle: Max Niemeyer Verlag.

Peperkamp, Sharon & Emmanuel Dupoux (2003). Reinterpreting loanword adaptations: the role of perception. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona: 367–370.

Polivanov, Evgenij Dmitrievič (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4:79–96 [English translation: The subjective nature of the perceptions of language sounds. In E.D. Polivanov (1974) *Selected works: articles on general linguistics*. The Hague: Mouton. 223–237].

Prince, Alan & Paul Smolensky (1993[2004]). *Optimality Theory: constraint interaction in Generative Grammar*. London: Blackwell.

Rose, Yvan & Katherine Demuth (2006). Vowel epenthesis in loanword adaptation: representational and phonetic considerations *Lingua* 116:1112–1139.

Rositzke, Harry A. (1943). The articulation of final stops in General American Speech. *American Speech* 18:39–42.

Schmidt, Anna Marie (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *Journal of the Acoustical Society of America* 99:3201–3211.

Schütz, Albert J. (1978). English loanwords in Fijian. In: Albert J. Schütz (eds.) *Fijian language studies: borrowing and pidginization. Bulletin of the Fiji Museum* 4:1–50.

Shinohara, Shigeko (1997). *Analyse phonologique de l'adaptation japonaise de mots étrangers*. Ph.D. dissertation, Université de Paris III.

Shinohara, Shigeko (2004). Emergence of Universal Grammar in foreign word adaptations. In: René Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in phonological acquisition*. Cambridge: Cambridge University Press, 292–320.

Shinohara, Shigeko (2006). Perceptual effects in final cluster reduction patterns. *Lingua* 116:1046–1078.

Silverman, Daniel (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9:289–328.

Smith, Jennifer L. (2006). Loan phonology is not all perception: evidence from Japanese loan doublets. In: Timothy J. Vance & Kimberly A. Jones (eds.) *Japanese/Korean Linguistics 14*. Stanford: CSLI, 63–74.

Steriade, Donca (2001). Directional asymmetries in place assimilation: a perceptual account. In: Elizabeth Hume & Keith Johnson (eds.) *The role of speech perception in phonology*. San Diego (CA): Academic Press, 219–250.

Tesar, Bruce (1997). An iterative strategy for learning metrical stress in Optimality Theory. In: Elizabeth Hughes, Mary Hughes & Annabel Greenhill (eds.) *Proceedings of the 21st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla, 615–626.

Tesar, Bruce & Paul Smolensky (2000). *Learnability in Optimality Theory*. Cambridge, Mass.: MIT.

Tsukada, Kimiko, David Birdsong, Ellen Bialystok, Molly Mack, Hyekyung Sung & James Flege (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics* 33:263–290.

Uffmann, Christian (2006). Epenthetic vowel quality in loanwords: empirical and formal issues. *Lingua* 116:1079–1111.

Vennemann, Theo (1988). *Preference laws for syllable structure and the explanation of sound change: with special reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.

Weiand, Klara (2007). Implementing Escudero's model for the subset problem. *Rutgers Optimality Archive* 913.

Yang, Byunggon (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24:245–261.

Yip, Moira (1993). Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2:261–291.

Yip, Moira (2006). The symbiosis between perception and grammar in loanword phonology. *Lingua* 116:950–975.

Perception, production and acoustic inputs in loanword phonology*

Andrea Calabrese

1. Introduction

This work started as a purely linguistic investigation of the phonological adaptations found in loanwords. However, the importance that speech perception plays in accounting for them soon became apparent, and a number of interesting observations resultantly came to light. Accordingly, this paper has become a study of both loanword phonology and speech perception. Central to this topic is the issue of how one perceives and learns unfamiliar sound configurations, or better, words containing unfamiliar sound configurations, and how these sound configurations are adjusted during this process.

As we will see, loanwords are generated by bilinguals when they take words from one of the languages they know and use them in another of the languages they know. In this case the adjustments that the loanwords undergo occur during speech production. However, there is still another situation in which loanwords are produced: monolinguals may learn new words from a language they don't know, or know poorly, to fill a lexical gap in their language. In the present study I will focus on loanwords generated in this way. The issues of speech perception, and of word learning, are fundamental in this latter case, and will therefore be central in the analysis in this paper.

Generally speaking, the goal of speech perception is the determination of the meaning of an utterance that generated a given acoustic input. This is achieved by identifying the words present in the utterance, and establishing their syntactic organization. When we want to learn foreign words, or even new words in our own language, however, the main goal of speech perception is the identification of their phonological shape so that we can properly memorize them. This involves

---

constructing—by means of an inferential computation— a mental representation of the words in terms of articulatory features, including fully specified syllabic and prosodic structures.

I will argue that identifying, or perhaps more aptly, "understanding", a sound, or a combination of sounds, means identifying the "instructions", i.e. the feature configurations,[1] characterizing its articulation in the production component. When a learner is faced with a new language, he must deal with sounds that are not included in the inventory of his language. The featural configurations characterizing these segments or combinations of segments are therefore absent in the production system, and the foreign sounds cannot be articulated.[2] It follows that

---

1.   See later in this section on features as "instructions".

2.   In Calabrese (1995, 2005) this means that there are active marking statements forbidding these feature configurations/combinations.

In this paper I will try to avoid using any particular formal theory of phonology. However, my own way of seeing phonology (see Calabrese (2005)) will surely influence some of my theoretical choices. To make those theoretical choices more clear, I will give a brief synopsis of the theoretical beliefs relevant to this work in this footnote.

The phonological system of a language is a historically determined complex set of output phonological representations derived from mnemonic representations by phonological operations. The input and output representations of the derivation must be such that they are able to interface properly with the relevant physical/mental component. Therefore, output representations must be able to be properly articulated by the motor system and properly perceived by the sensory system. Input representations must be such that they can be encoded in long-term representations in the memory. The proper interface properties of output representations, i.e. their ability to be pronounced and perceived, are determined by the constraints and rules contained in the markedness module. These constraints and rules trigger operations that convert illicit configurations into licit configurations that can be interpreted by the sensory-motor system.

The markedness module includes universal negative constraints such as the *prohibitions* and *marking statements*. Prohibitions identify configurations that are never possible for articulatory and/or acoustic/perceptual reasons. An example is *[+high, +low] which is necessarily articulatorily impossible insofar as the tongue body cannot be raised and lowered simultaneously.

Marking statements identify phonologically complex configurations that may be found in some but not all phonological inventories. An example is *[−back, +round] which "marks" the feature configuration [−back, +round] as phonologically complex. Another is *Complex Onsets which "marks" complex onsets as difficult. The reasons for their complexity or difficulty are due to independent properties of the sensory motor system that are reflected in the grammar through these constraints (see below). Marking statements may be active or deactivated. If a marking statement is active in a language, the complexity of this configuration is not accepted in this language, and segments containing this configuration are absent from the language.

the learner cannot "understand" them when they are heard. He lacks a mental representation of them. Through auditory exposure and motor training, the learner can learn to produce them, and therefore "understand" them in the perception process. However, such learning is difficult and time consuming. The other way is to adjust their featural representations to make them familiar, and therefore the learner can "understand" them in perceptual mental representations. These adjustments are implemented through the same repair operations used in the production system. The perceptual mental representation that is so obtained is then stored in long term memory, and becomes the adjusted underlying representation for the non-native sound configuration.

I assume that this process is common to all experiences of non-native sounds, in particular learning a new language or borrowing words with foreign sounds or sound combinations from a different language. In the latter case, given that there is no need to preserve the phonological and morphological shape of the foreign word, as in language acquisition, the foreign word can fully undergo adjustments that can be both phonological and morphological.

If perception involves interpretation and computation, as discussed above, it loses its primary function of tracking external reality, the environment; it becomes detached from reality and prone to illusions. Although illusion-like, interpretative failures may occur, as discussed in later sections, I assume that listeners always have a direct access to the acoustic signal. I will further assume that a representation of the acoustic signal is stored in a short-term acoustic memory ("echoic

---

If a marking statement is active in a language, the configuration marked by this statement is not accepted in this language. This configuration is *illicit*. Illicit configurations are fixed up by set of repairs provided by UG.

In this text when a segment or another structure cannot be articulated because it is "foreign" this implies the existence of marking statements characterizing the segment or structure in question as illicit; therefore it requires repair.

Marking statements and prohibitions belong to the grammar; they are grammatical statements about phonological representations. However, they are also interface conditions, i.e. the means through which the linguistic computational system is able to interpret and read the properties of the sensory-motor system. The markedness constraints represent the sensory-motor system in the linguistic computational system.

In particular, active marking statements indicate the absence, or unavailability, of computational programs converting phonological representations into articulatory ones. When a marking statement becomes active, the targeted phonological configurations cannot be transformed into articulatory commands; the repair procedures that occur in this case must then refer to the manipulations of phonological configurations that make this transformation possible (see Calabrese (2005) for more discussion of this model).

memory" see Neisser (1967)). I will propose that there is also an echoic long term memory, and argue that the acoustic representations preserved in echoic memory tie perception to external reality.

We will see evidence for two perceptual systems. One is dedicated to picking up environmental information, bottom-up; the other one is the reverse, functioning top-down, and is dedicated to analysis, identification, and recognition. The first system is tuned to the environment. In the case of speech, it collects and stores the acoustic signal—echoic memory is part of it. This system implements the acoustic analysis of the signal extracting its invariant spectral properties. It also discriminates new, unfamiliar sounds and sound strings from familiar, previously heard ones. What is new is sent for further analysis to the second perceptual system while it is temporarily stored for possible comparisons. The second system is then active in analyzing those new or unfamiliar configurations. The production component plays a basic role in it insofar as this system analyzes linguistic material by synthesizing it anew—it is an analysis-by-synthesis system (Halle & Stevens 1962). This system is fundamental in learning insofar as analysis of what is new is crucial for learning.

As in Calabrese (2005), here I assume a realistic approach to language (Bromberger & Halle 1992, 1997, 2000; Halle 2002), according to which. "phonology is about concrete mental events and states that occur in real time, real space, have causes, have effects, [and] are finite in number." (Bromberger & Halle 2000:21). If we look at production, this realistic view of language assumes that phonological theory investigates the system of knowledge that allows concrete occurrences of real time computational steps that convert mnemonic representations of utterances into articulatory representations. This knowledge involves representations and computations that have concrete spatio-temporal occurrences allowing for the production of concrete articulatory events which stem from the workings of an actual brain with all its limitations.

When we turn to perception, especially if we interpret the term "realism", naively, such an approach should lead us to focus on the concrete reality of the linguistic signal that is perceived. This reality is acoustic. Perception should then simply involve **bottom-up** processes that extract all the relevant perceptual information from the acoustic input as it comes in, without recourse, or with minimal recourse, to **top-down** processes involving independent linguistic knowledge. All of the information needed for the identification of the words and morphemes contained in an utterance should be present in the acoustic signal according to this view.

Evidence shows that this concept cannot be maintained. Problems with this idea are brought to light by Liberman (1957). One very striking finding in his

research was that, due to coarticulation, acoustic cues for consonants especially are highly context sensitive. For example, take the syllables /di/ and /du/. The information critical to the identification of these syllables is the transition of the second formant. However, that transition is high in frequency and rising in /di/, but low and falling in /du/. In the context of the rest of each syllable, the consonants sound alike to listeners. Separated from context, they sound different, and they sound the way they should sound: two "chirps," one high in pitch and one lower. Acoustically, they do not have a plausible common denominator, an invariant property, despite the fact that they are perceived as the same sound. Liberman recognized that, despite the context sensitivity of the acoustic signals for /di/ and /du/, naturally produced syllables do have one thing in common: they are produced in the same way. Therefore, articulatorily, the consonants have a single common denominator. Both syllables are produced by using the tongue tip to make a constriction in the alveolar region. Listeners' percepts are based on the articulatory reality of sounds.

Facts such this show that the identification of sounds in acoustic inputs needs information that is not immanently contained in these inputs, and that therefore cannot be simply extracted **bottom-up** from them, rather must be computed—**top-down**—through processes that can access the production system.[3] Notice that this does not contradict the realistic approach to language proposed above which assumes that linguistics deals with concrete mental events and states. In this view, both bottom-up and top-down processes involve concrete mental events and states that are an organic part of the perceptual experience.

Top-down processes in perception are also needed for other reasons. We know that listeners may "restore" missing phonetic segments in words (Samuel 1981; Warren 1970), and talkers shadowing someone else's speech may "fluently restore" mispronounced words to their correct forms (see Marslen-Wilson & Welsh 1978). This ability to restore missing phonemes or correct erroneous ones can be explained only if we assume top-down processes that access information in the lexical entries. Even more significant departures of perceptual experience from the stimulus may be observed in some mishearings (for example "popping really slow" heard as "prodigal son" (Browman 1980; Fowler 1986)) or "mow his own lawn" heard as "blow his own horn" (Garnes & Bond 1980; Fowler 1986). As for mishearings, Garnes and Bond (1980) argue that "active hypothesizing on the part of the listener concerning the intended message is certainly part of the

---

3.   The computational module that performs these operation was called the *phonetic module* by Liberman and Mattingly (1985).

speech perception process. No other explanation is possible for misperceptions which quite radically restructure the message … " (p. 238).

Access to the production system, to lexical information, and the ability to reconstruct, or even construct phrases and sentences in the percept indicate that perception clearly involves top-down processes. In this work, as mentioned previously, I propose that a top-down system plays a fundamental role in speech perception. The speech production component is part of this top-down system. This is the system that implements the analysis-by-synthesis of linguistic inputs.

One could propose, as in the motor theory of perception (Liberman & Mattingly 1985; Liberman & Whalen 2000), that perception is essentially a top-down process of construction and interpretation where bottom-up processes accessing the acoustic input have a minimal role. In this article I will also argue against this hypothesis, and point out a fundamental problem faced by top-down perception: if perception is essentially analysis through production, we should expect that sounds that cannot be articulated in production, e.g. foreign sounds, should also be unperceivable. Therefore, a learner could never be able to access them and learn them. This is contrary to the common human experience of foreign sounds. A learner can hear a foreign sound even though he cannot articulate and recognize them, and will try to learn them—"apprehend" them, as proposed later—by constructing articulatory representations that approximate their acoustic reality. The acoustic reality of the speech input must also be accessible in perception. To account for this fact, I will propose that a fundamental role in perception is played by echoic memory where acoustic images, that is, acoustic representations of inputs, are stored. As proposed above, echoic memory is part of the bottom-up component of perception where a preliminary acoustic analysis of the speech input is implemented. Thus following a model such as Trace (McClelland &Elman 1986; see also Klatt 1980), but in a strictly computational formulation, I will argue that perception must contain both a bottom-up and a top-down component that run in parallel and interact with each other.

Finally, a fundamental assumption of the present paper which must be made explicit before concluding the introduction is that in long term memory morphemes and words are represented as sequences of discrete segments each of which is characterized by a bundle of distinctive features. There is overwhelming phonological evidence that this is the correct interpretation, though I will not discuss this evidence here, but refer to Kenstowicz (1994), Halle (2002) and others for the compelling phonological arguments supporting this view. It is also important to highlight that acoustic evidence also supports such a view. Acoustic studies (Stevens 1972, 1989, 2002) of sounds produced by various manipulations of the vocal tract show that certain distinctive and stable acoustic patterns occur when the vocal tract is in particular configurations or performs particular maneuvers—these

configurations or maneuvers correspond to distinctive features. As Stevens (2002) points out, these combinations of acoustic and articulatory patterns are based on the physics of sound generation in the vocal tract, including theories of coupled resonators, the influence of vocal-tract walls on sound generation, and disconti-nuities or stabilities in the behavior of sound sources. Evidence for features also comes from quantal aspects of auditory responses to sound, such as responses to acoustic discontinuities and to closely-spaced spectral prominences (Chistovich & Lublinskaya 1979; Delgutte & Kiang 1984; Stevens 2002).

Distinctive features have a dual function. First, they serve as mnemonic devices that distinguish one phoneme from another in speakers' memories, a function that is fundamental during speech perception. Each feature also serves as an instruc-tion for a specific action of one of the movable parts of the vocal tract, a function that is fundamental during speech production (cf. Halle 2002).

The phonetic substratum for each feature establishes a link between a specific articulatory action and an acoustic and perceptual consequence of this action. As proposed by Liberman and Mattingly (1985, 1989), Halle (2002), Halle and Stevens (1991), the computational system that makes them capable of acquir-ing command of one or more languages includes a module, which I will assume is part of the top-down perceptual component, that selects specific actions of the articulators and links them to selected aspects of their acoustic consequences (Halle &Stevens 1991). These correlations between articulatory activity and acous-tic signal are controlled by the distinctive features. For example, the forward and backward placement of the tongue body is correlated with specific differences in the frequency of the second formant – this correlation is controlled by the feature [back]. Other examples include the correlations of the different placements of the tongue blade, be it before or behind the alveolar ridge, with the differences in the acoustic spectrum between hissing and hushing sounds, a difference controlled by the feature [anterior]. Similar relations between articulatory activity and acoustic signal are provided for each of the roughly nineteen features that comprise the universal set of phonetic features (Halle 1992; Halle & Stevens 1991).

In the next section, I will briefly review the most recent theoretical models in loanword phonology. After this brief review, I will discuss the perception model proposed here and demonstrate that this model offers the most adequate account for the adaptations of the foreign sounds found in loanwords.

2.   Loanwords

I begin by considering the nature of loanwords. First of all, one can distinguish two types of loanwords: *integrated loanwords* and *on-line adaptations* (Peperkamp

2002). I*ntegrated loanwords* are words that have entered the lexicon of the borrowing language. Monolingual speakers who use these loanwords never hear their source forms, and so the phonological analysis of the modifications these words have undergone when entering the borrowing language has no direct psychological reality. Rather, it receives a diachronic interpretation, in that it accounts for the adaptations applied by those speakers who originally introduced the loans. The *on-line adaptations* are foreign words that are borrowed 'here-and-now' (see, for instance, Shinohara (1997, 2000) and Kenstowicz & Sohn (2001)). In this paper, following Peperkamp, I treat integrated loanwords and on-line adaptations on a par, assuming that the former reflect on-line adaptations by those speakers who once introduced these words.[4]

Consider now the conditions under which linguistic borrowing occurs, i.e. the conditions that lead to the formation of loanwords. Assume two languages L1 and L2. L1 is the borrowing language and L2 the loaning language. Borrowing occurs when a speaker of L1 "borrows" a word of L2 to fill a lexical gap in L1. The reasons for this lexical gap can be many: lexical or cultural innovation may introduce objects or actions that do not have a name in L1; certain words may be felt as non-prestigious; certain words may simply be unknown, or just forgotten; new words may be created for playing, etc.

In any case, there are two possible scenarios in which borrowing occurs:

I.   A speaker is bilingual in L1 and L2. A lexical gap in L1 is filled in by taking a word from L2. The speaker retrieves the underlying representation of this word from his L2 mental dictionary (the long-term memory storage for L2

---

4.   In the case of the integrated loanwords, we also need to distinguish loanwords that underwent morphological nativization like the Italian loanwords (ia) from those that did not (b):

(i) a.   bistecca (bistecc+a)   from   English *beef-steak*
           birra (birr+a)   German *bier*
           giulebbe (giulebb+e)   Arabic   *gandulāb*
     b.   sport
           jeep [dʒip]
           killer

The morphologically nativized loanwords in (a) are characterized by the addition of affixes and by other morphological and phonological changes characteristic of the native grammar. Often it is difficult to distinguish these words from those that etymologically belong to the native lexicon. These loanwords can be treated with the other integrated loanwords, but they also require an analysis of the processes of morphological nativization that applied to them. I will not deal with this type of processes here (see Repetti 2003, 2006, This volume, for an analysis of changes of this type in the English loans to Italian.)

lexical items) and generates its surface representation while speaking in L1. If the surface representation of the word is generated by using the phonological, or more generally the grammatical, system of L2, the word is pronounced as in L2. There are no adjustments or adaptations. However, if the surface representation of the word is generated by using the phonological, or more generally the grammatical, system of L1, the word undergoes adaptations and adjustments. It is nativized according to the L1 grammar.

Examples of this type of borrowing can be found in utterance below, directions that were given to me in the Boston North End by an Italian American shopkeeper when I asked him where I could park my car to get to his store, a notoriously difficult enterprise. We had known each other for a long time. His response to my question, which I asked in Italian, appears below:

(1)   *Gira al corno di quella stritta. Poi prendi la seconda stritta a destra e vai stretto per due blocchi. Puoi parcare il carro proprio li.*

(2)   Lexical Borrowings from English into Italian

English	Loanword	Italian counterpart
corner	corno	angolo
street	stritta	strada
straight	stretto	diritto
blocks	blocchi	——
park	parkare	parcheggiare
car	carro	macchina

He was born in the Campania region of Italy, and had come to the USA when he was a teenager. He spoke Italian pretty well, although in his regional accent.

In the example above, we can see that my friend replaced the appropriate Italian roots, which may have slipped his mind while speaking, with their English counterparts. The English roots were adjusted by modifying their phonology (in terms of changing vowel quality and the gemination of word-final obstruents) and by adding Italian suffixal morphology. These adjustments were obviously done on-line while he was producing his utterance.

II.   The speaker of L1 does not know L2 well. He fills a lexical gap in L1 by learning the relevant word from a L2 speaker. Once the learned word will be uttered publicly or even silently, it becomes a loanword. Given that the speaker does not speak L2 well, the word will display adjustments and adaptations. There are two possible hypothesis to account for this. The first is that during perception and learning, the acoustic representations of the non-native segments is faithfully mapped into abstract featural representations (Jacobs & Gussenhoven (2000)). This featural representation is then modified

during production. The second is that the modifications already occur during perception and learning. Below we will see evidence supporting the second hypothesis.

Therefore, adjustments and adaptations may occur during speech production, and during speech perception and analysis, and the situation appears to be identical in both cases. There is, however, an obvious difference in the inputs. In the first case, the input to the adaptations and adjustment is an abstract long-term memory representation, an underlying representation. In the second case, the input is the acoustic signal produced by the surface phonetic representation of the word. This difference, obviously, determines the final shape of the loanword. We will consider the consequences of this difference in the next section.[5]

3. The current major accounts of non-native phonological adaptations

There are two models of loanword phonology: one assumes that borrowing occurs only in scenario I; the other assumes that borrowing occurs only in scenario II.

3.1   The phonological repair model

This model proposes that nativization is brought about by the phonological processes characterizing speech production. Thus, LaCharité and Paradis (2005) (see also Paradis & Tremblay (this volume) and Jacobs & Gussenhoven (2000) as well as the analysis of loanwords in Calabrese (1988, 1995) and Connelly (1992)) attempt

---

5.   It is important to note that during the process of adaptation, the adapter, exposed to a foreign sound from a language X, may also opt to acquire it as-is. Thus this sound may appear in the loanwords from language X. The same may occur for foreign syllabic and prosodic structures. For example, in Itelmen (J. Bobaljik (p.c.)), voiced stops occur only in loanwords. This type of innovation leads to situations in which the loanwords form X have a different phonology—different segments and possibly also other phonological and morphological phenomena derived from X—from that of the native words. They will thus form a special lexical stratum or co-phonology (Ito & Mester (1999). See Bobaljik (2006) for arguments in favor of this approach to loanword phonology.

Bobaljik (pc.) also referred me to Boberg (1991) who shows that for many speakers of American English there is a general phoneme "foreign A" [a:] which is used for foreign words like "pasta", "Mazda", "llama", "spa", "tobacco" regardless of source, but which are distinct from any other particular phoneme of (native) American English. As Bobaljik observes, in this case it is not surface phonetic similarity to the target that predicts the nativization pattern in the case of these words, just identification of them as belonging to a special [+foreign] class of words, and then regional and sociological factors.

I will not discuss cases of this type here.

to build production-based repairs into their nativization model. They assume that adapters start with underlying representations containing the non-native segments because the adapters are bilingual (LaCharité & Paradis 2005). Repairs to these non-native segments are implemented so as to avoid the production of marked or illicit segments or strings. A characteristic feature of these approaches is that speakers adapt loanwords by operating on a phonological/phonemic level that abstracts away from the details of allophonic and phonetic realization. The input to the adaptations is an abstract morphophonemic representation of the L2 word. For example, LaCharité and Paradis (2005) discuss the adaptation of English loans into French where the English lax high vowels /I/ and /U/ are mapped to French /i/ and /u/ instead of the acoustically closer /e/ and /o/. If loanwords are adapted in terms of operations on distinctive features then the configuration [+high, −ATR] of English /I/ and /U/ can be repaired into the configuration [+high, +ATR] of the French high vowels /i/ and /u/ regardless of the fact that French mid /e/ and /o/ are better acoustic matches in surface phonetic representations

Adaptations occur only during speech production in this model (cf. scenario I above).

3.2   Acoustic approximation model

According to this model, the adaptations we observe in loanwords are based on phonetic approximation/similarity. This model was first proposed by Hermann Paul in 1880. In his discussion of loanword phonology, he hypothesizes that a host speaker, upon encountering a foreign segment, matches this phonetic signal with the native segment with which it is most closely related. Paul implicitly assumes that this match involves a perceptual similarity judgment based on *Sprachgefühl*, the feeling of language: speakers adapt a non-native segment to one which they 'feel' most closely resembles the former acoustically.

New models of loanword phonology that acknowledge the importance of perception in determining similarity as the basis for the treatment of the loanwords (Silverman 1992; Yip 1993; Kenstowicz 2001, 2003 (see also Hsieh, Kenstowicz & Mou (this volume)) go back to this nineteenth century framework. According to them, the replacement operation between the non-native segment and the native one is strictly based on phonetic similarity between the outputs of the donor and recipient languages. Thus, according to Peperkamp and Dupoux (2003), the equivalences in loanword adaptation are based on a similarity that is defined as "acoustic proximity or proximity in the sense of fine-grained articulatory gestures."

In this model, the input to the adaptations is a surface phonetic representation of the L2 word and the similarity judgments producing the adaptations occur only during perception (cf. scenario II above)

3.3   Ito, Kang and Kenstowicz (2006)

Ito, Kang and Kenstowicz (2006) demonstrate that both models fail to account for the adaptation of Japanese vowels in Korean. The standard phonemic vowel inventories of the two languages are given in (3):

(3)
i	u		i	ɨ	u
e	o		e, ö	ʌ	o
	a		ɛ		a
Japanese			Korean		

The Japanese vowels are a subset of the Korean inventory. If we assume the acoustic/perceptual model of nativization, given their different sizes, the vowels of the two systems might be expected to partition the articulatory- acoustic space differently. However, when we examine the loanword correspondences, we find that most vowels pick out exactly the phonologically matching Korean vowels (see (4)).

(4)   Japanese   Korean

beNtoo	pent*oo	'boxed lunch'
azi	aci	'horse mackerel'
hako	hak*o	'box'
kagami	kakami	'mirror'
sebiro	sepiro	'suit'
teNpura	tenp*ura, temp*ura	'tempura'

There is one systematic exception. Japanese /u/ is adapted with the Korean central vowel /ɨ/ when it appears after the coronal sibilants [ts], [s], and [(d)z]. Elsewhere, it is adapted as Korean /u/ as shown in (5).

(5)   Japanese   Korean

a.
baNgumi	paŋkumi	'program'
unagi	unaki	'eel'
jurumi	jurumi	'relaxed'
gaku	kak*u	'frame'

b.
[ts]umi	s*ɨmi	'stack, pile'
jaki[ts]uke	jak*is*ɨk*e	'glazing, baking'
ko[ts]uzai	kos*ɨcai	'iron bar'
susi	sɨsi	'sushi'
suimono	sɨimono	'type of soup'
mizuage	micɨake	'unloading catch of fish'
kazunoko	kacɨnok*o	'herring roe'

This apparent exception is easily explained. Japanese /u/, frequently transcribed as the unrounded high back vowel [ɯ], is realized as centralized [ɨ] after [ts], [s], and [z] (Homma 1973:352–3; Fitzgerald 1996). This is an example where a

nondistinctive variant in the source language coincides with a phoneme in the borrowing language (Iverson & Lee 2004).

Ito, Kang and Kenstowicz argue that although the Japanese /u/ lacks lip round-ing, it is articulated with vertical lip compression (Vance 1987; Ladefoged & Maddieson 1996; Okada 1999). In other words, the Japanese high back vowel is produced with narrowing of the lip opening but without lip protrusion. The allo-phonic realization of the laryngeal fricative /h/ as a bilabial fricative [ɸ] before /u/ is indicative of this labial component. Here, to account for the these facts, I propose a modification of the feature [round], splitting it into two different features [labial] and [lip protrusion] where [+labial] is implemented by the narrowing of the lip opening. These two features are related by the marking statement in (6a) and the prohibition in (6b):[6]

(6)   a. *[+labial, −lip protrusion]
      b. **[−labial, +lip protrusion]

(6b) characterizes the configuration [−labial, +lip protrusion] as being always impossible. (6a), though, is characterized by high complexity and is therefore it is rarely deactivated (although there are some exceptions, like Swedish and Japanese). Thus, vowels with lip compression, like [+labial], are usually produced with lip protrusion, i.e. they are [+round]. However (6a) is deactivated in Swedish. Accordingly, a contrast between front rounded vowels with [+lip protrusion] and [−lip protrusion] can be observed in this language as in (7) (from Ladefoged & Maddieson (1996)):

(7) | +labial (lip compression) | +labial (lip compression) | −labial |
     |---|---|---|
     | +lip protrusion | −lip protrusion | −lip protrusion |
     | ryːta | ruːta | riːta |
     | 'roar' | 'window pane' | 'draw' |

(6a) is deactivated in Japanese, but not in Korean. The feature specifications of the relevant vowels are given in (8):

(8)   a.   Japanese   i   ɨ   ɯ
           [back]   −   +   +
           [labial]   −   −   +
           [lip protrusion] −   −   −

      b.   Korean   i   ɨ   u
           [back]   −   +   +
           [labial]   −   −   +
           [lip protrusion] −   −   +

---

6.   See Note 2 for a brief discussion of the notions of marking statement, prohibition and repairs in the model of phonology adopted here.

Therefore, in borrowing Japanese words, Koreans repairs the illicit configuration [+labial, −lip protrusion] of Japanese by delinking [−lip protrusion] and inserting the unmarked [+lip protrusion], as in (9), where Lip is the articulator node from which the two features [labial] and [lip protrusion] are dependent:

(9)   /ɯ/   [u]

          Lip   →   Lip   →   Lip

      +labial   −lip protrusion   +labial   +labial   +lip protrusion
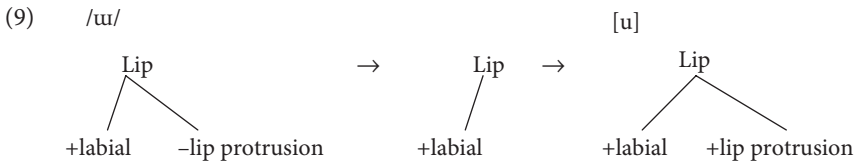
In the case of the post-sibilant [ɨ] allophone of Japanese, we must assume that it is [−labial] as shown in (8).[7] Therefore this allophone is featurally identical to the Korean [ɨ].

   Ito, Kang and Kenstowicz show that the Korean adaptations of Japanese vowels are problematic for the two models of loanword adaptation discussed above. The fact that the Korean adaptation takes account of the Japanese [ɨ] allophone is problematic for the phonological model of LaCharité and Paradis (2005). This segment would not be expected to be present at the phonemic level, where the loanword phonology operates in their model. Yet it is precisely in the post-sibilant context that Japanese /ɯ/ is adapted as Korean /ɨ/, strongly suggesting that the adaptation is taking account of this predictable allophone.[8]

   A model of loanword adaptation that assumes that mapping is strictly based on phonetic similarity between the outputs of the donor and recipient languages (Silverman 1992; cf. Peperkamp 2002; Peperkamp & Dupoux 2003) would also fail to provide a straightforward account for the adaptation of Japanese /ɯ/. When we examine the acoustic properties of these vowels, given the close proximity of the Japanese /i/, /e/, /a/, and /o/ and their Korean counterparts in acoustic space, the adaptation of these vowels can be accounted for in terms of acoustic similarity. However, this model incorrectly predicts that Japanese [ɯ] should be adapted as Korean [ɨ], and not [u], since they are the most similar in the acoustic map, as illustrated in (10) (figures from Ito, Kang & Kenstowicz (2007)).

---

7.   Evidence for this is that in the Yonaguni dialect of Okinawa (Joo 1977, p. 125) /u/ after /s/, /z/, /t/ and /d/ has merged with /i/ (see Ito, Kang & Kenstowicz (2006)).

8.   Note that /u/ occurs without problems after the sibilants in Korean (e.g. *supak* 'watermelon'), thus precluding an independently motivated adjustment /u/ → [i] after sibilants.

(10)
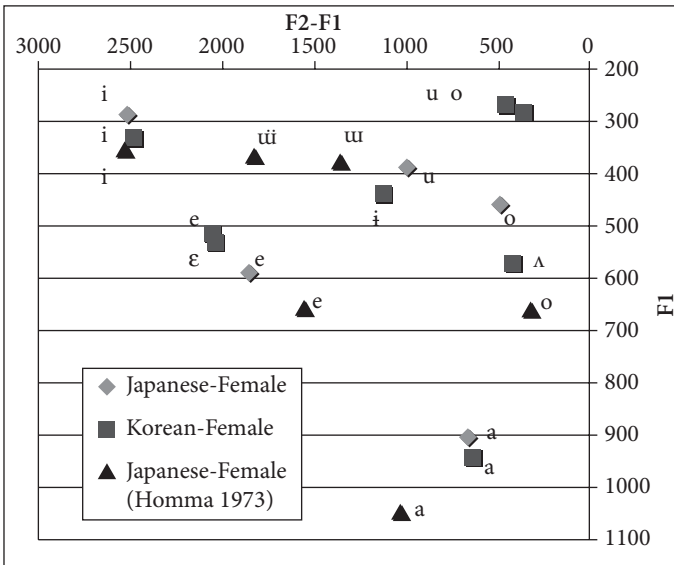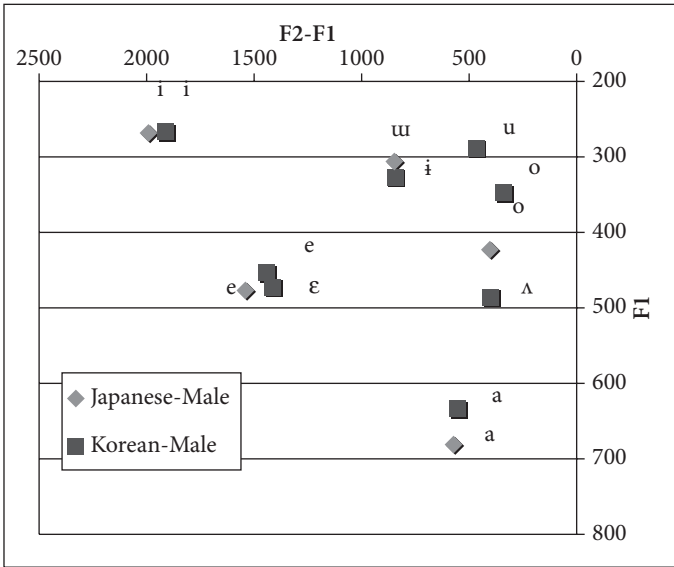
No matter how the various formants of the high vowels are weighted, as Ito, Kang and Kenstowicz observe, Japanese [ɯ] is best matched by Korean /ɨ/ in purely acoustic terms. Nevertheless, in loanword adaptation Japanese [ɯ] is adapted as Korean /u/ except after sibilants.

In light of these observations, Ito, Kang and Kenstowicz argue for a third model of loanword adaptation: This model assumes that enough phonetic detail

must be retrieved from the donor language so as to distinguish the two allophonic variants of Japanese /ɯ/. It follows that the input to the adaptation must be a surface phonetic representation, At the same time, this model also assumes that the adaptations operate on abstract featural representations of the source language: they involve involve phonological operations on features.

4.   Evidence for loanword adaptations in perception

Peperkamp and Dupoux (2003) review psycholinguistic evidence showing that all aspects of non-native phonological structure, including segments, prosodic, and syllable phonotactics, are systematically distorted during speech perception; i.e. non-native sound structures are adjusted both by monolinguals and by bilinguals. Comparing loanword adaptations to experimental speech perception data, they point to a number of striking correspondences. For instance, Korean listeners find it hard to distinguish between the English consonants [r] and [l] in CV stimuli (Ingram & See-Gyoon 1998), and in English loanwords word-initial [l] is adapted as [r] (Kenstowicz & Sohn 2001). In a similar vein, French listeners have severe difficulties perceiving stress contrasts (Dupoux et al. 1997) and in loanwords, stress is systematically word-final, regardless of the position of stress in the source word.

A striking case which demonstrates the fundamental role played by speech perception in the nativization of loanwords is the perception of illusory vowels in consonant clusters by Japanese speakers. These individuals perceive illusory epenthetic vowels in sequences of segments that do not fit the syllable structure of their native language. Moreover, Japanese speakers often insert epenthetic vowels when they pronounce loanwords involving these same clusters:

(11)   a.   [ma.kɯ.do.na.rɯ.do]   'MacDonald' (Japanese)
        b.   [sɯ.to.ra.i.ko]   'strike' (Japanese)

At this point, an obvious question arises: Are such epenthetic vowels inserted in production or perception? In a series of behavioral experiments, Dupoux and colleagues (Dupoux et al. 1999; Dupoux et al. 2001; Dehaene-Lambertz et al. 2000) compared Japanese listeners with French listeners in their perception of consonant clusters. For instance, Dupoux et al. (1999), give an off-line phoneme detection task (Experiment 1) in which they used a series of six items created from naturally produced nonce words (e.g. [abuno], [akumo], [ebuzo], [egudo], etc.) in which they gradually reduced the duration of the vowel [u] to zero milliseconds. While listening to a recording of the sounds, participants were asked if the item they heard contained the sound [u]. Japanese listeners, unlike French listeners, overwhelmingly

judged that the vowel was present at all levels of vowel length. Strikingly, this was the case seventy percent of the time even when the vowel had been completely removed. The French participants, on the other hand, judged that the vowel was absent in the no-vowel condition about 90% of the time and that a vowel was present only in 50% of the intermediate cases. These results were confirmed in other experiments, which have led Dupoux and colleagues to conclude that the influence of native language phonotactics can be so robust that listeners perceive illusory vowels to accommodate illicit sequences of segments in their L1.

Similarly, Kabak and Idsardi (2006) show that Korean listeners perceive illusory vowels within consonantal clusters that are illicit in that syllable structure. As was observed for Japanese, Korean speakers also insert epenthetic vowels within the same clusters. (cf. [a.i.sɯ.kʰɯ.rim] 'ice cream', [khɯ.ri.sɯ.ma.sɯ] 'Christmas').

Given the overall similarity between speech perception data and loanword adaptations, Peperkamp and Dupoux (2003) propose that all loanword adaptations apply in perception. Though this is too strong, loanwords with adaptations can appear in situations of bilingualism as in the North End/Italian example discussed previously. However, it seems that many or even most borrowings do not occur in a situation of bilingualism, but in a situation of language contact between monolingual or imperfectly bilingual speakers. Here the role of perception and learning is fundamental. Observe that, as discussed above, for loanwords that develop in this situation, the input is a surface phonetic representation. In this paper I will focus on such borrowings.

Peperkamp and Dupoux (*ibid*) also assume that the adaptations observed in perception involve phonetically minimal transformations. That being said, it is unclear how adaptations such as epenthesis or stress shifts, are phonetically minimal. Furthermore, the adaptations discussed by Ito, Kang and Kenstowicz which have surface phonetic representation as inputs, and therefore must have occurred in perception, are not phonetically minimal, but actually appear to involve clear phonological operations.

In this paper, I will follow Peperkamp and Dupoux (*ibid*) in assuming that at least some, if not most, of the phonological adaptations characterizing loanword phonology occur during perception. Moreover, I will also propose that these adaptations involve the same phonological processes that characterize speech production.[9]

Before going on, however, I would like to address a fundamental problem of a philosophical nature: if perception involves production, or better construction of a representation through production, we would be experiencing only fallible

---

9.   See Kim (this volume) and Boerma and Hamann (this volume) for a similar approach to loanword phonology.

representations – illusions. We would be detached from the hardness of reality. This cannot be correct. We do experience reality directly, or in the case of speech, we do have a direct access to the acoustic reality of the message. I will try to address this issue in the next section.

5.   Perception

As first pointed out by Immanuel Kant in the 18th century, perception is not neutral, or passive, but rather involves a complex inferential computation by which sensory data from a given object are organized, categorized and adjusted by accessing abstract cognitive categories and previous knowledge.

At the same time, perception must be anchored to reality. As observed by Gibson (1979) and Merleau-Ponty (1945), perception must be tied to our being in the environment. Individuals must also have a direct, unmediated experience of their surroundings so that they can interact with it properly. Perception through inference is obviously indirect and prone to mistakes and does not capture the concrete experience of the world obtained through perceiving bodies immersed in the moving texture of reality.

Following Norman (2001), I will propose that these two aspects of perception correspond to two different but interacting perceptual systems: the "ventral" system dedicated to the identification and recognition of objects and events in the environment and the "dorsal" system dedicated to picking up information for a proper interaction with the reality. In the first system perception is indirect, mediated by the cognitive system and memory; in the second system it is direct, immediately given in the process of sensory-motor integration.

Recent work in the cortical organization of vision has emphasized that sensory input must interface both with conceptual systems (for object recognition) and with motor systems (e.g. visually guided reaching/grasping) (Ungerleider & Mishkin1982; Milner & Goodale 1995).

It has been demonstrated empirically that these two interface systems comprise functionally and anatomically differentiated processing streams in which a ventral (occipital-temporal) stream supports object recognition/understanding (the "what" pathway), and a dorsal (occipital-parietal) stream supports visual-motor integration functions (the "where" pathway).

The primary function of the ventral system is the recognition and identification of objects and events in one's environment (Norman 2001). It compares visual inputs to stored information in an attempt to achieve a meaningful interpretation of those inputs. The ventral system deals mainly with the utilization of visual information for interpreting one's environment, The recognition and identification

processes that are part of this interpretation require comparisons with stored representations. This system is therefore memory-based, using stored representations to recognize and identify objects and events.

In contrast, the primary function of the dorsal system is the analysis of visual inputs from the ambient array in order to allow interaction with the environment and the objects in it e.g. pointing, reaching, grasping, walking towards or through something, climbing, etc. (Norman 2001). The dorsal system picks up visual information to allow the organism to function in the environment. It is a system that picks up invariants in the ambient array by directly "resonating"—to use Gibson's (1979) ecological terminology — to the features to that array. It is a system that has a direct relationship with the environment. It anchors the perceiving individual to the external reality. Some have suggested that all the information pick up for the performance of well-ingrained actions or behaviors is implemented by the dorsal system. The dorsal visual stream is thus particularly geared for visual-motor integration, as required in visually guided reaching and orienting responses (Rizzolati, Fogassi & Gallese 1997).

The dorsal system deals mainly with the utilization of visual information for the guidance of behavior in one's environment. Much of our day-to-day pick up of visual information is carried out by the dorsal system without much conscious awareness. As Norman (2001) points out, this system is ecological in the sense of Gibson (1979). It directly picks up information from the ambient array for or through action. From now on, I will refer to the outputs of this perceptual system as "*sensory intuitions*" insofar as they have a unmediated, direct, or "naive" relation to objective reality as in Kantian empirical intuitions.

The ventral system is instead, a "higher" system that deals with the interface between the visual input and cognition, and we are normally conscious of its output. It is the system that tries to make sense of situations. This is achieved via an indirect, post-sensory, inferential nature. Interpretation is an inextricable part of the perception processes characteristic of the ventral system. This is the system that implements the type of indirect, inferential perception advocated by constructivist models of perception (see Helmholtz 1867; Rock 1977, 1983, 1997; Gregory 1993; Boring 1946; Epstein 1982; see also Piaget (1937, 1969); Marr (1982); Anderson (1985); Kanisza (1979).[10] I will refer to the outputs of this perceptual process with

---

10.　In this way, following Van Leuween (1998) and Norman (2001), the Gibsonian approach on perception, in which perception is simply the pick up of information from invariants in the ambient environment, and which by itself, is inadequate as a general model of perception, can be integrated with a more adequate constructivist model of perception.

the term "*apprehensions*". In the terminology adopted here, visual perception thus includes both "sensory intuitions" and "apprehensions."

6.   Speech perception

Following Hickock and Poeppel (2004), I assume that two similar perceptual systems are also active in speech perception. But in order not to create confusion between visual perception and speech perception, this text will refer to the linguistic ventral system **as the top-down system**, and the dorsal system as the **bottom-up system**.

The objective of listeners during the act of speech perception is to access the meaning of the utterance they hear, this can be accomplished only via the identification of the vocabulary items (morphemes/words) used in it and the interpretation of their structural significance in the syntactic environment in which they occur. Crucially this identification/recognition depends on the identification of the sounds comprising the vocabulary items. In classical semiological terms, one must identify the *significans* (the form of the expression) in the utterance to access its *significatum*.

Given that identification and recognition are characteristic functions of the top-down (ventral) system, one could propose that speech perception is implemented only by the top-down system. However I propose that the bottom-up (dorsal) system also has a role insofar as it picks up information and provides it to the top-down system for analysis. Thus, the bottom-up system picks up acoustic data and converts them into acoustic representations —"sensory intuitions"— that are provided to the top-down system which converts them into articulatory representations — a process of "apprehension". These articulatory representations can be used to identify and recognize the vocabulary items, which, as discussed in Section 1, are represented in terms of articulatory features. Thus, whereas the dorsal system in vision functions as a parallel perceptual system dedicated to visual-motor integration, in the case of speech perception where the main goal is identification and recognition of *significanta* in the acoustic input, a goal that can be achieved only by the top-down system, the bottom- up system must pick up acoustic information from the input and produce "sensory intuitions" or the top-down system where these representations are interpreted/apprehended.[11]

---

11.   But note that later, following Hickock and Poeppel (2004) I will propose that there is a direct path between acoustic representations in echoic memory and meaning/concepts, therefore a speech perception stream that by-passes the top-down system. This path between

Following Blumstein and Stevens (1981), Stevens and Blumstein (1981), and Stevens (1998), I assume that the acoustic signal is analyzed to extract invariant acoustic properties or patterns which excite designated acoustic detectors in the bottom-up system. An acoustic representation is produced through this parsing of the acoustic signal. As proposed above, this retrieval of linguistic information from the acoustic signal is implemented by the bottom-up system.

The point is that in contrast to the discretely specified phonological representation of an utterance, the acoustic signal that is produced by a speaker is continuous. It is an analog signal that is generated by continuous movements of a set of articulatory and respiratory structures. However, as mentioned in Section 1, the relations between the articulatory and the acoustic representations of speech have certain quasi-discrete or quantal characteristics that are exploited in speech production (Stevens 1972, 1989). These quantal attributes help to simplify the process of uncovering the discretely specified segmental and categorical representations that are building blocks of words (Stevens 2002).

As proposed by Stevens (2002), the retrieval of linguistic information from the acoustic signal proceeds as follows. First, the locations and types of the basic acoustic landmarks in the signal are established. These acoustic landmarks are identified by the locations of low-frequency energy peaks, energy minima with no acoustic discontinuities, and particular types of abrupt acoustic events. From these acoustic landmarks certain articulatory events can be hypothesized: the speaker produced a maximum opening in the oral cavity, or a minimum opening without an abrupt acoustic discontinuity, or a narrowing in the oral cavity sufficient to create several types of acoustic discontinuity (Stevens 2002). Such landmarks provide evidence for distinctive features such as [consonant], [sonorant], [continuant], [strident], the so-called articulator-free features in Halle (1995). The second step consists of the extraction of acoustic cues from the signal in the vicinity of the landmarks. These cues are derived by first measuring the time course of certain acoustic parameters such as the frequencies of spectral peaks or spectrum amplitudes, in particular frequency ranges, and then specifying particular attributes of these parameter tracks (Stevens 2002). These acoustic parameters provide evidence that articulators are involved in producing the landmarks and demonstrate how these articulators are positioned and shaped.

However, coarticulation and other contextual and prosodic adjustments affect cues and landmarks reducing their strength or just eliminating them. Thus, uncovering of the segments and features that underlie the words in an utterance, then,

---

"acoustic" *significantia* and *significata*, however, is possible only under special circumstances, namely an established familiarity of the *significans*.

involves using acoustic data to make inferences about the gestures that the speaker uses to implement these features, since the gestures tend to bear a closer relation to the features than do the acoustic patterns, as first pointed out by Liberman (1956) and discussed in Section (1). This inferential activity is part of the top-down system. Stevens (2002) proposes that once the sequence of landmarks has been identified and a set of acoustic cues has been evaluated in the vicinity of each landmark, the next step is to convert this landmark/cue pattern into a symbolic or quantal description consisting of a sequence of feature bundles. This conversion is carried out in a specialized phonological module, which, according to my proposal, forms part of the top-down system. This module consists of a set of submodules, one for each feature. Every submodule examines the acoustic landmarks and cues that are relevant to the feature and analyzes them in term of their overall environment and prosodic context. It then produces an inference concerning the specification of the feature. The details of each of these submodules, including the selection and design of acoustic cues that form possible inputs for each module, are beyond the scope of this article (see Stevens (2002) for more discussion).

Through the activity of the phonological module, the acoustic signal of an utterance is converted into an array of articulatory features. I assume here that each submodule interprets the acoustic data autonomously from the other submodules. It is the duty of a special component called the synthesis component, to read out all of the articulatory information provided by these submodules to construct an underlying representation which is necessary to access the meaning of the utterance and then derive from it the surface representation which is behind the utterance that produced the acoustic input. Recognition/identification/apprehension of the acoustic input consists of this process of interpretation and construction (see Section 10 for some hypotheses on how this occurs). One of the first steps in the synthesis component, after the piecing together of the information provided by the phonetic module, is the identification of the morphemes and words present in them by matching them with morphemes and words that are stored in a long term memory dictionary as arrays of features. This matching process is then followed by a top-down construction of a sentence which attempts to make sense of the morphemes and words that have been activated in long-term memory. I propose that speech perception thus also involves activation of the production component of the grammar that implements the computation involved in this construction (see Section 10 for more detail).

I assume that the mapping into articulatory representations, the analysis and top-down construction of the sentence discussed above are implemented in the verbal working memory, the so called phonological loop (Baddelley 1992). It is here that the analysis-by-synthesis of the sentence is implemented. Hickok and Poeppel (2004) argue that verbal working memory relies on a auditory–motor integration

network (Aboitiz & García 1997). Indeed, verbal working memory (and perhaps working memory in general) can be viewed as a form of sensory– motor integration (Wilson 2001). For example, in Baddeley's model (1992), the phonological loop is essentially a mechanism for using motor systems (articulatory rehearsal) to keep sensory-based representations (the phonological store) active. Hickok and Poeppel postulate an explicit neural network for verbal working memory. In their model, the phonological store is identified with the Superior Temporal Gyrus systems support-ing acoustically-based representations of speech which the articulatory rehearsal component maps onto frontal systems supporting articulatorily-based represen-tations of speech. Evidence from Awh et al. (1996) indicates that the articulatory rehearsal component involves left frontal cortices, notably portions of Broca's area and more dorsal pre-motor regions.

Notice that a working memory phonological buffer (the phonological loop of Baddeley (1992)) is fundamental to understanding language acquisition (Doupe & Kuhl 1999): in fact, for the child to learn to articulate the speech sounds in his or her linguistic environment, there must exist the following components: (i) a mechanism by which sensory representations of speech uttered by others can be stored (echoic memory); (ii) a mechanism by which the sensory input is analyzed and converted into an articulatory representation (the analysis-by-synthesis component of the pho-nological buffer (see Section 10), (iii) a mechanism by which the child's articula-tory attempts can be compared against these stored representations, and by which the degree of mismatch revealed by this comparison can be used to shape future articulatory attempts (the "comparator" in the top-down system (see Section 10)). Although such a network obviously assumes less importance in adult speakers, the fact that new words from one's own language and from foreign languages can be learned at any age show that it continues to operate throughout life.[12]

We expect the motor system to be active in this working memory compo-nent. And in fact, recent neurological studies have shown that perceiving speech involves neural activity of the mirror neurons and the motor system. The mirror neurons are a particular class of neurons that exhibit excitations not only when an individual executes a particular action but also when the same individual observes the action being executed by another individual. The existence of these neurons (see Di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolati 1992) provides direct neural evidence for motor system involvement in perception. (Rizzolati & Craighero 2004;

---

12. Further evidence that the phonological loop plays a role in adults is provided from articulatory decline following late-onset deafness (Waldstein, 1989), from the effects of delayed auditory feedback on speech articulation (Yates, 1963), and from altered speech feedback experiments (Houde & Jordan, 1998).

Rizzolati, Fogassi & Gallese 2001). In Rizzolati and Arbib's (1998) words, "taken together, the human and monkey data indicate that, in primates, there is a fundamental mechanism for action recognition. ... Individuals recognize actions made by others because the neural pattern elicited in their premotor areas during action observation is similar to that internally generated to produce that action." (p. 190)[13]

Two recent studies involving the use of transcranial magnetic stimulation of the motor cortex have demonstrated activation of speech-related muscles during speech perception. Fadiga et al. (2002) found that when listeners hear utterances that include lingual consonants, they show enhanced muscle activity in the tongue. Watkins and colleagues (Watkins, Strafella & Paus 2003) found that both while listening to speech and while seeing speech-related lip movements, people show enhanced muscle activity in the lips. Complimentarily, two fMRI studies (Pulvermüller et al. 2006; Wilson, Saygin, Sereno & Iacoboni 2004) demonstrated that there is overlap between the cortical areas during speech production and those active during passive listening to speech. As shown by Fadiga et al. (2002), the same motor centers in the brain are active both in the production of speech and in speech perception, where the perceiver engages in no overt motor activity.[14, 15]

---

13.   Recently researchers have shown that Rizzolati and Arbib's (1998) "fundamental mechanism for action recognition" also has ties with general audition. Kohler et al. (2002) found neurons in the pre-motor cortex of monkeys that respond not only when the monkey performs a specific action (e.g. breaking a nut) or sees the action performed by someone else, but also when the monkey merely hears the sound that is caused by the specific action (e.g. the cracking noise of the nutshell being broken).

14.   Still, as discussed below, it is possible to have basic speech perception, in the sense of access to meaning of an utterance, without accessing the top-down system and the motor systems.

15.   Fadiga, along the lines of the motor theory of speech perception proposed by Liberman and Mattingly (1985, 1989), or better in terms of the Direct Realism model of Fowler (1986, 1994, 1996), states that "speech perception and speech production processes use a common repertoire of motor primitives that during speech production are at the basis of articulatory gesture generation, while during speech perception are activated in the listener as the result of an acoustically evoked motor 'resonance'" (Fadiga et al. 2002). Fadiga, therefore, assumes that there is direct perceptual relation between acoustic stimuli and motor activity, which he calls an acoustically evoked motor 'resonance', as in the ecological theory of perception of Gibson (1979). But how this resonance is implemented is unclear. Infact, at the neural activity of the motor and premotor systems does not need to be seen in terms of Gibsonian 'resonance' but could be seen rather in a more indirect, memory-mediated manner in which these neural activations are brought about by the phonological loop, as discussed above.

7.   Phonological adaptations in perception

In the normal situations of speech interaction, the goal of speech perception is comprehension, the identification of the *significatum*, the meaning carried by the *significans* of the utterance. Identification of the *significans* is fundamental in the learning of new words regardless of whether the new word belongs to the learner's native language. In either case, a mental representation of the *significans* of the word to be committed to memory must be constructed. This representation must include both phonological and morphological information. In order to be committed to memory, however, the *significans* must be interpretable, i.e. identified and recognized by the cognitive systems Therefore, if it contains unfamiliar or illicit, configurations, i.e. uninterpretable configurations, they must be adjusted so to be identified.

When faced with an unfamiliar linguistic sound, a perceiver has an obvious problem insofar as a configuration that is uninterpretable in terms of his own system of linguistic knowledge must be analyzed in terms of this system. A first rough account of what happens in this case is the following. If a segment, or a syllabic combination of segments, is unfamiliar, foreign, i.e. absent fom L1, a speaker has no instructions for how to produce it, i.e. no representation of it with the right combinations of features, or segments in the case of syllable configurations. In particular, as proposed in Calabrese (2005) (see also Footnote 2), the absence of a segment, or a syllabic combination of segments, indicate the absence, or unavailability, of computational programs coordinating their featural configurations into articulatory commands. The absence of such a program is formalized as an active constraint against the configurations. Configurations violating an active constraint are repaired. The repair that occur in this case then indicate the featural/configurational manipulations that adjust the representations and make their conversion into articulatory commands possible (see Calabrese (2005) for more discussion of this model). As proposed in Calabrese (2005), active constraints are checked throughout the derivation, and if violated, repairs apply.

If production plays a role in perception, then it follows that active constraints and repairs play a role as well. Accordingly, unfamiliar sounds disallowed by active constraints must be repaired in perceptual representations, thereby resulting in a perceptual adaptation of the unfamiliar sounds.

At this point it is useful to focus a hypothetical instance of this process. The event begins with the production of an utterance by a foreigner in his or her own language. The native listener does not speak this language well. Therefore, this utterance contains sound configurations that are novel to him/her and are disallowed by an active constraint in his grammar. This utterance produces a certain acoustic signal, which enters the bottom-up perceptual system and is recorded in

the acoustic short-term memory storage, the short-term echoic memory. In the acoustic analysis component, this signal is analyzed in terms of its invariant acoustic properties, as discussed in the preceding section, but it is still an uninterpreted "sensory intuition". Thereafter, this acoustic information is sent to the top-down perceptual system for identification/recognition. The phonological module converts it into articulatory information by converting the acoustic invariant properties present in the acoustic data into articulatory features/structures. Remember that each submodule of the phonological module interprets the acoustic data autonomously from the other submodules and that the synthesis component of the phonological working memory buffer uses the pieces of articulatory information (feature specifications, prosodic and syllabic articulatory cues, etc.) provided by the submodules to construct an underlying representation and then derive from it a surface representation. If the acoustic input contained new or unfamiliar sounds or sound configurations, the synthesis component cannot put together the pieces and cues provided by the phonological module in the case of these sounds, or sound configurations. In fact, if these pieces and cues were put together as indicated in the acoustic input, they would form illicit, "unpronounceable" configurations blocked by active constraints of the grammar. These illicit configurations must then be adjusted and repaired. The application of these repair operations will produce a "more familiar" nativized representation.[16]

Consider a concrete example. Take an Italian listener like me. The Italian vowel system does not have the [+low, −back] vowel /æ/ of the English word /kæt/. In the terminology developed above, this means that there is an active constraint disallowing the feature configuration [+low, −back]. The acoustic input of this vowel will be characterized by a low first formant and a higher second formant. The submodule for the feature [low] detects the first acoustic property and assigns the specification [+low] to the articulatory representation of the word /kæt/ that is being built in the synthesis component. The submodule for the feature [back] assigns to it the feature specification [−back]. The information provided by the acoustic input through the phonological module also requires the simultaneous articulation of these two features in the vowel. However, the synthesis component operates according to the grammar of L1. Thus the fact that there is an active constraint disallowing this feature combination in L1 prevents the synthesis component to put together these two feature specifications into the same feature bundle. Thus, this feature combination cannot be created in the mapping from the acoustic invariant properties of the signal into the articulatory representation of the synthesis

---

16. See Kim (this volume) for a model in which acoustic cues for L2 sounds are mapped into L1 featural representations without repair operations.

component insofar as it is illicit in Italian. This featural configuration must then be repaired; it may either by changing the feature [+low] into [−low], deriving [ɛ] from /æ/, or by changing the feature [−back] into [+back], deriving [a] from /æ/. It follows that I may interpret/apprehend /kæt/ as [kɛt] or as [kat]. Accordingly, the acoustic configuration characterizing this vowel (with its low first formant and higher second formant) cannot be associated with the appropriate articulatory feature configuration ([+low, −back]), resulting in a "phonetic illusion".[17]

---

17.   The input to the analysis and interpretation in linguistic perception is the word in its surface representation. This is how the L2 word is "heard". Assuming a general principle of economy of operations is executed during perception, just as in production, as proposed in Calabrese (2005), only the configurations of the L2 word that are illicit in L1 must be repaired, i.e. and not those that are licit in L1. Only the minimum that is necessary to fix the input is changed, the rest must be preserved. The preservation of the licit aspects of the input may explain why the treatment of L2 words in L1 often involve processes that are not part of L1 phonology. One such discrepancy between processes in native and loanword phonology is provided by the treatment of English loanwords into Korean. In Korean, [s] is not allowed in syllable codas. In the native phonology, an underlying /s/ is realized as [t] when it occurs in coda position (ia). However, in English loanwords, words with [s] in coda position systematically undergo epenthesis (ib) (Kenstowicz & Sohn 2001).

(i)   a.   /nas/ [nat] 'sickle-NOM'
              /nas + ɨl/   [nasɨl]
       b.   [posɨ]   <   'boss'
              [kɨrasɨ]   <   'glass'
              [mausɨ]   <   'mouse'
              [kʰarisɨma]   <   'charisma'

The reason for the selection of epenthesis instead of neutralization to remove the coda consonant results from the desire to preserve the featural composition of [s] that is otherwise licit in the language (see Boersma & Hamann (this volume) for an account of these facts in a framework similar to the one proposed here but adopting OT).

       However, not all cases of a discrepancy between processes in native and loanword phonology can be accounted for in a similar way. Consider, for illustration, the treatment of loanwords from French into Fula. In the latter language, neither onset nor coda clusters are permitted. In loanwords from French, an epenthetic vowel is added after the second consonant in liquid+obstruent clusters see (iiai), but between the consonants of obstruent+liquid clusters see (iiaii) (Paradis & LaCharité 1997). In the native phonology, however, the epenthetic vowel is always inserted after the second consonant, both in the case of liquid+obstruent clusters, as in (iibi), and in the much rarer case of obstruent+liquid clusters, see (iibii) (data from Paradis (1992)).

(ii)   a.   i.   [karda]   <   Fr. carde [kard] 'card (comb)'
                     [fɔrsɔ]   <   Fr. force [fɔrs] 'force'
              ii. [kalaːs]   <   Fr. classe [klas] 'flag'

At this juncture it is important to deal with a common aspect of the speaker/listener's experience of foreign sound that is aptly described in the following anecdote (from Jacobs & Gussenhoven (2000: 203)): "Valdman (1973) reports the reaction by a Haitian-Creole-speaking maid who attended evening literacy classes to her teacher's pronunciation of oeuf 'egg' as [ze]: she decided to leave the class. Although she herself pronounced it that way, she was aware that her bilingual employers realized it as [zø]". In other words, she did not know how to pronounce the word, but she did know how it sounded. A listener can be aware of the acoustic shape of a given sound, although his own pronunciation is different.[18]

The simplest hypothesis to account for a situation like the one above is that upon presentation of a foreign word/sounds, a learner/listener learns perfectly faithful representations of them without access to the production system which we know is impaired in lacking relevant articulatory instructions for the foreign sounds. These faithful underlying representations would be modified by the speakers when they are pronounced during production; at the same time they could be used as criteria to judge the difference between their correct forms in L2 and their adapted forms in the speaker pronunciation, as in the above event. This would require a perceptual system that is independent of the grammatical system which faithfully converts acoustic inputs into underlying featural representations (Jacobs & Gussenhoven (200); Hale & Reiss (2000)).

The evidence provided earlier that the perceptual representations of listeners are distorted by their native grammar (Dupoux et al. 1999; Dupoux, et al. 2001; Dehaene-Lambertz et al. 2000) shows that this hypothesis cannot be maintained. Learning the exponent of a new or foreign word means learning the articulatory patterns expressed in its featural composition. If there are grammatical constraints against the featural organization of these patterns, the word cannot be learned as such and its featural organization must be adjusted, as discussed above. Therefore

---

b.   i.   /talk+ru/   →   [talkuru] 'amulet'
      ii.   /sokl+ka/   →   [soklaka] 'need'

It is unclear to me how preservation of the licit featural configurations of the input can account for the difference between native and loanword phonology in this case. Such cases are perhaps better analyzed by hypothesizing a special loanword phonology grammatical component. This component would include the active marking statements of the native language and the related repair processes that were grammaticalized/istitutionalized during the contact with the foreign language. Further research is required to establish if such a special grammatical component is needed.

18.   The same occurs in children. When my daughter was about two year old, she used to pronounce [ʃ] as [s]; therefore, she said [sIp] instead of [ʃIp] 'ship'. So once I tested her and, while pointing to the picture of a ship, I asked: "Is this a [sIp]?". She replied, "No! it is a [sIp]!"

it follows that a faithful conversion of a foreign word into a long-term memory underlying representation is impossible.

Furthermore, Section 9 will address evidence that children older than 5 or 6 and adults are "behaviorally deafened" to foreign contrasts not shared with the native language. This further weakens the idea that it is possible to convert a foreign linguistic input into a faithful featural representation. It is true, as will soon be evident, that by focusing attention on the acoustic input and intensive training, this "deafness" can be overcome. But this training involves articulatory exercises that teach the learner to acquire the articulatory pattern that are constrained in his L1 grammar. Once these articulatory patterns are acquired, the L2 sound is learned and its accurate "perception" becomes possible. But this is precisely what is argued here. An accurate perception of a sound requires learning how to produce that sound articulatorily.

However, if what is perceived is simply identical to what is produced – as can be concluded at first from the above proposal, awareness of the acoustic shape of the given sound by a listener, when his own pronunciation of this sound is different, should be impossible. A more careful consideration of the theory shows that this consequence is incorrect. The bottom-up system in fact must include an echoic short-term memory (Neisser 1967) where aural acoustic representations[19] of speech sounds are preserved. These are the acoustic representations that are analyzed in the preliminary phase of speech perception.[20] We can also assume that such representations can be stored in long- term echoic memory with all other non-linguistic sounds.[21] Observe that, in the preliminary phase of speech analysis, where acoustic patterns are analyzed in terms of their invariant properties, there must be a basic ability to distinguish sounds so as to extract their characterizing features. It follows that we must be able to distinguish aural representations

---

19. The term *image* is often used in this case "aural acoustic image". I prefer to use representation to indicate that there is always a degree of symbolic conversion between the sound in itself and the sensory *representation* of it provided by our neural networks.

20. The output of this analysis must already be somewhat abstract insofar as linguistically irrelevant properties, such as the voice characteristics of the speaker that uttered the word, the rate of speech, distortions caused by a cold or sore throat and so on, can be neglected in the formation of the memorized linguistic representation of the sound. Obviously an issue here is how we store words in memory, be they as word-type or word-tokens (Goldstone & Kersten 2003; Hintzman 1986; Goldinger 1986, 1988). This issue falls outside the aegis of this article.

21. Anyone that has ever played with a short-wave radio knows that it is possible to understand if the language of the tuned station is Chinese, Russian, Arabic, etc. without knowing these languages – this is possible by accessing an aural memory of how those languages are spoken.

of sounds preserved in long-term echoic memory, and use them for comparison with items stored in the short-term echoic memory. It is therefore in the echoic long term memory of the Haitian maid that the standard sound [ø] for [zø] 'egg' was stored. Thus, although she could not pronounce it, she could use that echoic memory representation for comparisons.

Neurolinguistic evidence suggests that echoic memory is located in the auditory cortex (Näätänen 2001). If this is correct, one expects evidence of analytic activity in this part of the cortex though the fact is that several studies suggest that more complex events such as stream segregation – extracting the abstract sound patterns and invariant sound relationships – and categorical speech perception guided by language-specific memory traces may take place preattentively in the auditory cortex (Sussman et al. 1998, 1999; Tervaniemi et al. 1994; Paavilainen et al. 1999, 2001; Dehaene-Lambertz 1997; Phillips et al. 2000; Phillips 2001; Shestakova et al. 2002).

Distinguishing aural images of sounds, however, does not mean being able to recognize or identify them. Recognition and identification of a string of sounds involves determining how the feature specifications detected in the preliminary parse of the utterance acoustic input are organized into feature bundles and how these bundles are distributed in a syllabic string, i.e. it means identifying how these feature specifications are combined in the featural configurations of the string. To put this in Kantian terminology, we can say that only at this point could we have perceptual judgments and thus state whether or not a certain sound is /ø/ or /æ/, i.e. a feature bundle containing the configuration [−consonantal, −back, +round], or a feature bundle containing the configuration [−consonantal, +low, −back]. Insofar as the lack of relevant articulatory instructions block these feature configurations, the possibility of these judgments is prevented, and in so doing thus the recognition and identification of the relevant sounds is not possible.

It follows that a listener can be aware of distinctions among unfamiliar sounds without being able to identify them – this awareness is possible because of acoustic images of the sounds stored in echoic memory. However, although this listener may feel that these sounds are different, he cannot know how they are different, and when forced to represent them in the construction of the linguistic representation in apprehension, he must adjust them into more familiar configurations.

The problems pointed out by Ito, Kang and Kenstowicz (2006) are then easily addressed. The input to perception is a surface phonetic representation. If an allophonic property is licit and can therefore be interpreted it will appear in the loan. Thus Japanese back unround vowel [ɨ] found after sibilants, an allophone, can be adopted as such in Korean loanwords insofar as such a vowel is present in the Korean phonological system.

Furthermore, given that acoustic proximity does not play a role in the model proposed here, the Japanese back vowels with the configuration [+labial, −lip compression] will be repaired by Korean speakers as discussed in Section 3.3.

Japanese vowel epenthesis may be accounted for as follows: Japanese disallows complex onsets and complex codas. Simple codas are restricted to the first member of a geminate, or to a nasal glide. When presented with a word with such illicit clusters in the phonological working memory buffer, a Japanese speaker cannot construct a mental representation of it because of the active constraints active in his grammar. The listener resorts to epenthesis to fix this problem. The presence of the illusory vowels in perception then occurs.

Just as in the nativization of the high [−ATR] vowels of English to high [+ATR] vowels in loanwords to French, Spanish and Italian, I assume that we are dealing with a case of repair. The configuration [+high, −ATR] of English I, U is blocked by an active marking statement in these languages which is addressed by delinking the feature specification [−ATR] and replacing it with [+ATR]. This can occur both in perception or production as in all other cases discussed previously.

Following Dupoux et al. (1999), I used the term "phonetic illusion" to characterize the perceptual adjustments discussed above. But these are not "sensory illusions" in the sense that something is heard which is actually not part of the stimulus, rather, I assert that they are representational illusions which are generated by the linguistic operations that are used to construct adequate mental representations of linguistic objects. Given that these representations are memorized, the illusions are accessible only through recall. The Japanese listeners, in fact, report that there is a vowel in the consonantal clusters, but in that case they are "reporting" on the memorized representation of the stimulus, not on what they actually heard. To put this differently, they hear a certain acoustic input, say [eb—zo], as is but they cannot make sense of it perceptually because they lack this type of syllable structure in the production grammar. So they adjust the representation of this stimulus by epenthesizing a vowel [ebuzo]. It is on this memorized representation that they report, not on what they heard.

Once this is clear we have an account of the illusions found in the human experience of foreign speech sounds. Interestingly, these illusions are also observed in the experience of the native language, as most famously observed by Sapir (1933). While studying the Canadian Athabaskan language Sarcee, Sapir was puzzled by his informant John Whitney's insistence that there was a difference between *dini* 'this one' and *dini* 'it makes a sound' even though the two were phonetically homophonous to Sapir's trained ears. In order to explain his informant's intuition, Sapir postulated that the final vowels of words like *dini* make a sound with a "latent" consonant. Put differently, this suggests that there was another psychologically more accurate

representation of the word that records the presence of this intuited sound: [dinit]. As Sapir observed, this was the *phonological representation* of the word.

Sapir called this experience by his informant a *phonetic illusion*. Two objectively identical stimuli *dini* were judged as different when associated with different phonological representations. Additional examples of this type of illusions are easy to find. Kenstowicz (1994), for example, reports that English speakers tend to perceive the intersyllabic consonantal material in *camper* and *anchor* as analogous to *clamber* and *anger*. This is an illusion, however. In most dialects (Malecot 1960) the nasal consonant is phonetically absent before such sounds as [p,t,k,s], so that *camper* and *anchor* have the same gross phonetic shape (C)ṼCVC (Ṽ a nasal vowel) as (C)VCVC *wrapper* and *acre*. While ṼCVC *anchor* belongs with VCVC *acre* phonetically, English speakers have the strong intuition associating it with VCCVC *anger*. This perceptual judgment is likewise based on the abstract phonological representations of these words.

In the model proposed here, these illusions are accounted for by assuming that the perception process involves access to the abstract phonological representation computed in the mind.[22]

8. Echoic memory and sensory intuitions

The echoic memory in the auditory cortex stores the acoustic features of the stimulus (Neisser 1967). The sensory information stored by echoic memory covers

---

22. Another model that assumes that adaptation of sounds in perception is due to the access to the articulation system is the Perceptual Assimilation Model of Best (1994, 1995). Best treats the adaptations found in the pronunciation of foreign sounds as assimilations to the closest native phoneme category on the basis of articulatory similarities and discrepancies.

Best assumes that listeners perceive information about articulatory gestures in the speech signal. Thus, they can perceive in non-native phones information about their gestural similarities to native phonemes. If the listener perceives the phones to be very similar in their articulatory-gestural properties to a native phoneme category, then the nonnative phones will be assimilated to this native phoneme category.

Besides the fact that this model rejects the use of phonological features for which there is overwhelming evidence, as discussed in Section 1, it remains unclear how information about the similarity of articulatory gestures in native and non-native phones is actually extracted from the acoustic signal without detecting that there are also acoustic discrepancies. It seems that there is a fundamental circularity in this approach to loanword adaptations. Notice furthermore that without recourse to phonological features the notion of gestural similarity become quite arbitrary. What makes an articulatory gesture more similar to a certain gesture than to another gesture?

Unfortunately, lack of space prevents a deeper discussion of this model.

all aspects of the stimulus.[23] I assume that echoic memory allows direct access to the acoustic input, it stores what I call the sensory intuition of the acoustic message, and involves the sensory system that closely tracks the acoustic stimuli. Experiments of duplex perception provide evidence for this system. Mann and Libermann (1983) conduct one such experiment in which the components of a syllable were presented dichotically. The base, presented to one ear, included steady-state formants for /a/ preceded by F1 and F2 transitions consistent with either /d/ or /g/. An F3 transition, presented to the other ear, distinguished /da/ from /ga/. Perception is called duplex because the transitions are heard in two different ways simultaneously. Listeners perceive a clear /da/ or a clear /ga/ in the ear receiving the base depending on which transition was the other ear was exposed to. In the ear receiving the transition, listeners hear a non-linguistic chirp. The listeners could be asked to attend the syllable or the chirps and make quite different judgments on them. They responded quite differently, even though the judgments were based on the same acoustic input. On the one end, this can be seen as evidence of a special system interpreting the combined sensory input of both ears and producing a linguistic percept; this is the Top-Down system advocated here. On the other end, it provides evidence of a system tracking the acoustic signal and storing it as such, a system referred to here as echoic memory, part of the bottom-up system.

There is also evidence that the auditory cortex maintains more permanent representations of the auditory past. Recent neurological evidence suggest that there is a set of sound memory traces, and of memory traces of sound combinations representing syllables and words, of the language involved in the left auditory cortex (Näätänen 2001; Pulvermüller et al. 2001; Pulvermüller et al. 2006) When a familiar speech sound is presented, it activates the corresponding phonetic trace or recognition model (in addition to the different sound-analysis mechanisms common to speech sounds and equally complex non-speech sounds).[24]

I propose that learning to speak a new language, in addition to learning how to articulate the different combination of features and syllabic organization which characterizes that particular language, also involves the formation of discrimination and recognition patterns for the acoustic counterparts of these features. This acoustic discrimination and recognition patterns help the further analysis and processing of the acoustic signal into featural representations by the top-down system.

---

23.   Experimental evidence suggests that the memory-trace durations for echoic memory last between 5 and 10 seconds. Echoic memory can occur outside of conscious experience and attention-independently; (Woldorff, Hackley and Hillyard, 1991).

24.   This auditory cortex process itself is *pre-perceptual*, but tends to trigger frontal cortex activity which probably underlies the initiation of attention switch to sound change.

Hickock and Poeppel (2004), furthermore, provide evidence for a cortical network which performs a direct mapping between acoustic representations and conceptual-semantic representations. Learning words, and any sublexical process requiring the analysis of phonological exponents requires the mapping of acoustic invariant properties of the signal including these words into articulatory feature representations and subsequent processing in the top-down system. But, once a word is learned, used, and heard many times, becoming thus totally familiar, it is obviously uneconomical to go through the phonological module and analysis by the top-down system every time that the same word is heard. One could assume, instead, that recognition of familiar and commonly used words and constructions may simply bypass analysis by the top-down system, and directly activate the exponents of the dictionary in the phonological working memory buffer through a direct association between acoustic representations stored in long-term memory echoic memory and word exponents in the dictionary. Hickock and Poeppel demonstrate that this must be the case by providing a variety of evidence, the most compelling of which is the dissociations we observe in aphasia patients. Namely, Broca aphasiacs demonstrate an ability for word recognition and comprehension despite widespread damage to the production system, and to what is known here as the phonological working memory buffer. Comparatively, Wernicke aphasiacs and Word deafness syndrome patients are characterized by their inability to access the dictionary, and a subsequent lack of word comprehension; on the other hand, both have an intact production system, and an ability to repeat syllables and words, which is explainable only if we also assume an intact phonological module and phonological working memory buffer.

9.   Acoustic inputs and phonological discrimination

It is often stated that children become "deaf" to foreign phonological contrasts in the process of language acquisition.

Prior to the period of six to nine months of age, infants apparently can discriminate any sounds contrasts in any language. By the end of the first year, however, they apparently can no longer discriminate most sounds that do not contrast in the ambient language (Aslin, Jusczyk, & Pisoni 1998; Best, McRoberts, LaFleur, & Silver-Isenstadt 1995; Polka & Werker 1994; Werker & Lalonde 1989; Werker & Tees 1984). This developmental change is accounted for in language learning models (cf. Best's Perception assimilation model: Best 1994, 1995; Flege's speech learning model: Flege 1991; Kuhl's Native Language Magnet: Grieser & Kuhl 1989; Iverson & Kuhl 1996; Kuhl 1991, 1992; Kuhl et al. 1992) as a side effect of the infant having learned the phonological categories of the ambient language during this six to nine month period. It is proposed that by attracting both the ambient and foreign sounds the infant hears, these categories learned by 12 months deafen

him to differences detectable six months earlier when he had not yet learned any categories. This defeaning is temporary until apparently 5 to 6 years of age when, if the child does not get sustained exposure to the foreign language early enough, he will be permanently deafened behaviorally to foreign contrasts not shared with the native language. For example, adult speakers of languages with fixed stress (French, Finnish, Hungarian) are significantly less able to detect a shift in stress position in a word than a change in one of its segments (Dupoux, Pallier, Sebastian, & Mehler 1997; Dupoux, Peperkamp, & SebastiánGallés 2001; Peperkamp & Dupoux 2002). Even highly fluent bilinguals are deafened, too, if their exposure to the second language is too late. For example, some Spanish dominant Spanish-Catalan bilinguals who did not learn Catalan before 5–6 years of age cannot discriminate Catalan contrasts not shared with Spanish, high-mid versus low-mid vowels, /e, o/ versus /ɛ, ɔ/, or voiced versus voiceless fricatives, / z / versus / s / (Pallier, Bosch, & SebastiánGallés 1997).

If I am correct in assuming that acoustic differences can always be detected by the bottom-up system, the "deafening", or lack of discrimination mentioned above must be a consequence of the top-down system. The presence or lack of contrasts between sounds is governed by grammatical constraints (Calabrese 2005).[25] Consider the low-back vowel a/ and the front-low vowel /æ/. A language will contrasts these two sounds if and only if the constraint *[+low, −back] is non-active (see Calabrese (2005)). Comparatively, a language has only /a/, and therefore lacks this contrast, if this constraint is active. Constraints are obviously part of the grammar and therefore are also components of the top-down system. In Section 7, I discussed how active constraints trigger repairs that adjust non-native segments in the top-down perceptual system. These repairs lead to perceptual neutralizations of contrasts. When this occurs, two sounds cannot be recognized/identified

---

25.   I assume that learning a language involves learning which configurations are admissible. The hypothesis is that the child starts with an inability to produce all segments and combinations of segments except the basic, unmarked ones such as /a/, /m/, /t/, /ta/, /ma/, etc. (see Jakobson (1941)). Learning involves learning to produce the "marked" segments of the ambient language. If the approach proposed here is correct, the child must also be unable to recognize the "marked" sounds before s/he learns to produce them. However, s/he can hear them: s/he has a raw sensation of them in terms of the aural image of the signal present in echoic memory, as proposed earlier. Exposed to the featural configurations extracted from these "raw" stimuli, what I called the sensory intuitions, a child eventually constructs the appropriate combinations of articulatory features in the representations of words and vocabulary items and learns how to articulate them in production and to identify them in perception.

After the end of the critical period, the child loses the ability to easily learn to produce new phonological configurations. A possible way to look at this fact is by noting that the neural motor pathways become set after the critical period so that learning a new array of articulatory movements becomes difficult.

as phonologically/linguistically different. Phonological discrimination means recognition by the top-down systems that two sounds are linguistically/phonologically different. For example, given the case that I discuss in that section, in my own pronunciation of English, there is no contrast between vowels /a/ and /æ/ insofar as I adjust the latter in a different vowel; therefore the contrast between these two vowels is neutralized in my perception of English. I do not discriminate between them in the sense that I cannot recognize them as linguistically different. I can hear that they are acoustically different but I do not know how they are phonologically different, in the sense that I do no know how to account for their difference in articulatory terms, as discussed in Section 7.

Sounds that are recognized as linguistic by the top-down system obviously trigger special linguistic behavioral responses insofar as they are recognized as linguistically relevant in the construction of vocabulary items. Notice that acoustic properties characterizing the foreign sounds will be considered linguistically irrelevant by the linguistic attentional system, and neglected by it—listeners will not pay attention to them, and the acoustic discrepancies brought about by them will be difficult to detect linguistically, although they can be heard granted special attention. Therefore, as Kingston (2003) observes, the deafening observed in children does not mean an inability to hear differences between acoustic stimuli but rather refers to the weakening in the behavioral response to these differences. An infant is behaviorally deafened to foreign contrasts because he no longer responds differently when the stimulus changes from one member of a foreign contrast to the other. Although the infant can still hear the different acoustic stimuli, he just does not respond to them, as it would to native phonological contrast, insofar as they are linguistically unimportant.

For adults, deafening likewise does not imply an inability to hear or access the acoustic signal, but rather a lack of an ability to recognize acoustic configurations as phonological entities, and therefore to discriminate acoustic differences as instances of phonological contrasts. However these acoustic differences can be heard if the adult speaker is made aware of them and pays sufficient attention. If adults were really deaf to foreign sound categories, they would never be able to learn a second language. Indeed, with enough focus adults can hear the phonetic contrasts of a foreign language, and can try to learn how to produce them articulatorily.

This proposal is sufficient cause to warrant a reinterpretation of the Best's findings (1994, 1995). Best investigates how the adaptations of foreign sounds influence the listeners' ability to discriminate different foreign sounds from one another. The primary difference in adaptations is between assimilation of two foreign sounds to two versus just one native sound, "two category" (TC) versus "single category" (SC) assimilations, respectively. Best observes that listeners discriminate the members of TC assimilations far better than SC assimilations. For example, English listeners can discriminate Zulu lateral fricatives /ɬ - ɮ/ well, insofar as they assimilate these two sounds to two different segment categories. However, they do quite poorly with

the Thompson Salish ejective velar /k'/ and uvular /q'/ that are likely to assimilate to English [kʰ]. SC assimilations are further distinguished between those in which both foreign sounds assimilate equally to the single native category, as in the Zulu lateral fricative case, versus those in which the nonnative pair are both assimilated to a single native category, yet one may be more similar than the other to the native phoneme. This is the case of Zulu aspirated /kʰ/ and ejective /k'/: they both assimilate to English /k/ but / kʰ/ is more similar to the English surface allophone /kʰ/. In the latter case, according to Best, the two foreign sounds differ in "category goodness" (CG) with respect to the native category. The members of such CG assimilations are more easily distinguished than the other type of SC. Thus Best suggests that listeners' success in distinguishing different foreign sounds is ranked: TC > CG > SC. She also reports that listeners perform very well in discriminating sounds that cannot be easily assimilated to English sounds like the Zulu clicks.

Observe that all of these sounds are acoustically distinct for the listeners regardless of whether they are assimilated to native sounds or not assimilated (as with the clicks). Even in the case of the sounds entering a SC contrast, Best says that they are "heard as discrepant" by listeners (Best 1994:191). What changes is the listeners' capacity to interpret non-native acoustic discrepancies as phonological contrasts by processes of identification in the top-down system. Therefore, if the non-native sounds are identified by the top-down system as involving distinct native phonological categories, they can be discriminated as different. However, if the non-native sounds are identified by top-down interpretation and adjustments as involving a single native phonological category, i.e. if there is perceptual "neutralization", then discrimination is obviously impossible. In the CG, one of the sounds does not undergo any adjustment in the top-down system (as there is no active constraint against it) while the other does. This difference affects the perceptual process, and hence the actual perception of that sound.

10. A model of speech perception

Building on what is outlined above, I propose a speech perception model based on two components: (1) the assumption that both a bottom-up and top-down system are active in perception and (2) the idea that production in the top-down system has a fundamental part in perception of new words/utterances. To do this I turn to the analysis-by-synthesis model of speech recognition proposed by Halle and Stevens (1962) and adopted by Mattingly and Liberman in their motor theory of speech perception.

According to this analysis-by-synthesis model, the listener analyzes the acoustic input by deriving how it is produced by the speaker, synthesizes a virtual acoustic signal based on the output of this derivation, and matches the virtual

signal to the actual one. Given a sufficiently close match, the listener achieves a mental representation of the percept that corresponds to the invariant motor commands sent to the musculature underlying the vocal tract actions that produced the acoustic signal. The analysis-by-synthesis component is part of the top-down system. The complete model with the bottom-up and top-down components is schematically represented in (12):

(12)

TOP DOWN (INTERPRETATION/PRODUCTION) SYSTEM

Other Cognitive Modules

PERCEPTUM

16

15

4

SEMANTIC COMPONENT

ENCYCLOPEDIA
Extra-linguistic knowledge

DICTIONARY
Cohorts of VIs

CONCEPTUAL SYSTEM

SYNTACTIC/MORPHO-SYNTACTIC GRAMMAR COMPONENT

MORPHO-PHONOLOGICAL GRAMMAR COMPONENT

WORKING MEMORY
PHONOLOGICAL BUFFER

9

8

7

SYNTHESIS
OF HYPOTHESIZED REPRESENTATION

PHONOLOGICAL MODULE
Acoustic landmarks/cues > Articulatory features

ACOUSTIC ANALYSIS
It identifies pertinent landmarks and cues in acoustic signal

LONG TERM ECHOIC MEMORY

BOTTOM UP (INFORMATION PICK UP) SYSTEM

10

11

6

5

4

3

ARTICULATORY INTERFACE
Universal phonetics
(conversion of representation into articulatory commands)

COMPARATOR

14

12

13

ACOUSTIC MEMORY BUFFER

2

Only
In Production

17

OUTPUT
Muscular Implementations of articulatory commands

1

INPUT SIGNAL

The block diagram in (12) describes the architecture of the speech perception model proposed here. Italicized numbers in the text refer to the different arrows in (12).[26]

The input acoustic signal enters the bottom-up system and is placed in temporary memory storage pending completion of the analysis *(1)*.

The acoustic analysis in the bottom-up system identifies its invariant acoustic properties and provides a discrete decomposition of the acoustic signal into acoustic landmarks and cues *(2)*. This analysis is generated by a general acoustic module, not specifically dedicated to linguistic signals.

The representation that is so obtained is checked by the long term echoic memory storage system *(3)*. There are two possibilities at this point. (I) If this acoustic representation does not match an already stored representation, that is, it is new or unfamiliar, it needs to be "apprehended", i.e. identified/recognized by the top-down system and goes to step *(5)*. (II) If the acoustic representation matches an already stored representation, the latter becomes active. Once activated, the acoustic representation of familiar words/utterances in the echoic memory system on their turn directly activates the relevant conceptual structures *(4)*. The meaning of an utterance can therefore be directly accessed from the acoustic input in this case. This occurs when we are dealing with commonly used words, constructions and sentences which need to be analyzed only when they are first learned and perceived. Eventually the analysis becomes automatic, and their meaning is automatically associated with the acoustic representation generated in the acoustic analysis phase in the bottom-up system. A new analysis by the top-down system would therefore be unneeded and non-economical. In this case, perception simply bypasses the top-down system.

If the acoustic representation needs to be "apprehended" by the top-down system, it is first sent to the phonological module where the invariant acoustic properties (landmarks and cues) are interpreted as articulatory features *(5)*. For each feature there is a submodule that interprets the acoustic cues/landmarks of the signal and assigns a specification to the feature. Intonational, metrical and other prosodic cues should provide a segmentation of the utterance into phrases and words.

This surface representation of the utterance, which is "hypothetical" in so far as it is the outcome of inferences/interpretation by the phonological module, is sent to the working memory buffer. Here it is parsed/analyzed in the synthesis component in *(12)* and the production system of the working memory buffer *(7)*.

---

26. I currently consider the architecture in (12) and the relative discussion of this section speculative. It is a way of putting together my hypotheses on how language is produced/perceived and my ideas on how this system of perception/production is positioned in the mind/brain architecture.

A fundamental step in the analysis of this surface representation is the extraction of its underlying representation. This is done by checking cohorts of vocabulary item URs from the Dictionary against the hypothetical surface representation of the words provided by the phonological module. When a vocabulary item UR is chosen for the UR of a word, its morphosyntactic features and meaning is accessed in the dictionary and in the encyclopedia. These morphosyntactic features and meanings are checked against the morphosyntactic and semantic structures that are in the meanwhile generated in the synthesis component to account for the order of the elements and the phrasing in the representation of the utterance *((9)–(10))*. The meaning of the generated structures is also checked against the general pragmatic context. I assume that all of these processes/derivations occur in parallel and will eventually converge in producing a morphosyntactically and semantically well-formed surface sentential representation. Recognition/identification/"apprehension" occur at this point, namely, when a well-formed and licit representation is constructed from the inputs provided by the phonological module. To see if this process is successful, however, it is necessary to wait for the further steps in *(12)–(14)* when a virtual acoustic image of the representation is produced and then compared with the acoustic input stored in the echoic memory buffer, as discussed below.

In fact, the generated surface articulatory representation built from the inputs provided by the phonological module must be checked against the acoustic input to determine whether or not it is correct. This is done by generating a virtual acoustic synthesis of this articulatory representation. First the articulatory representation is converted into complex sets of articulatory commands/gestures *(12)*. These commands are then implemented silently without actual muscular activity, thereby creating a virtual acoustic synthesis that is sent to a comparator module *(13)*. In the comparator, virtual acoustic synthesis is checked against the acoustic input stored in the acoustic memory buffer *(14)*. If there is a successful match, the comparator instructs *(15)* the phonological working memory buffer to read out the representation whose phonological content and morphological, syntactic and semantic structure produced the match and to release it *(16)* as a perceptum to other cognitive modules *(17)*.

At this point it is necessary to consider how the checking and matching of VI URs from the dictionary against the inputs provided by the phonological module is implemented. For the sake of simplicity I consider only what happens in the case of a word simply composed of a root. The same must be done for all of the words and morphemes composing the utterance. There are several possibilities: (I) there is one UR that matches the featural configurations of the word provided by the phonological module; (II) There are more URs that have this perfect match (a case of homonymy); (III) There is a UR that matches the featural configurations

of the word provided by the phonological module only partly; or (IV) There are more URs that partly match these featural configurations. Consider possibility I and II first. In case I, there is no problem, the UR is chosen unless the morpho-syntactic or semantic context are incompatible with that choice. In case II it is the morphosyntactic and semantic context that determines the selection of the UR. In this case, other similar URs must be tried and the one that is compatible with the context is chosen. In the cases III and IV, it is necessary to access the phonological component *(11)*. The processes (rules and repairs) included in this component are applied to the URs in the relevant order to generate surface representations (see Chapter 14 of Anderson (1992)). The UR of the generated surface representation that matches the acoustic input of the word under analysis is chosen as the UR of this word unless this selection is incompatible with the context. Just as before, if there are more possible selections, the one compatible with the context is chosen. Again I assume that all of these derivations and processes run in parallel until a successful match is reached.

The idea that speech is perceived by reference to production assumes that in the perception process, in particular in what I call *apprehension*, the listener has an active role: he is able to access abstract morphological and syntactic levels of rep-resentation of the perceived utterance and compute its surface articulatory shape from these abstract levels. A successful perceptual act occurs when the acoustic shape of the articulatory representation derived in this perceptual computation matches the acoustic input in the auditory memory.

It is important to stress the fundamental role that the production system has in the model. Utterances are generated following the same steps as those discussed in the analysis above. The morphosyntactic component generates the hierarchical organization of the sentence *(10)* which is also computed by the semantic compo-nent *(9)*. In this hierarchical structure the UR of the vocabulary items stored in the Dictionary are inserted *(8)*. The surface representation is then derived by applying the phonological and morphophonological processes of the morphophonological component *(11)*. The crucial difference is that the articulatory commands orga-nized in the articulatory interface *(12)* are implemented in the muscular system, and therefore an actual acoustic signal is produced *(18)*.

Observe that in addition to *apprehension* where the listener has an active role, simultaneously a more passive perceptual process can also occur as in the case of commonly used words, constructions and sentences. Although such items are analyzed when they are first learned and perceived, the analysis becomes automatic, and their meaning is automatically associated with the acoustic rep-resentation generated in the acoustic analysis phase in the bottom-up system. As proposed above, in this case perception bypasses the phonetic module and working memory.

The model in (12) predicts both Top-down and Bottom-up effects in speech perception. These kinds of findings are often described as evidence for an inter-action of "bottom-up" and "top-down" processes in perception (e.g. Klatt. 1980). Bottom-up processes analyze the acoustic signal as it comes in. Top-down pro-cesses draw inferences concerning the signal based both on the fragmentary results of the continuing bottom-up processes and on stored knowledge of likely inputs. As discussed in the introduction, top-down processes can restore miss-ing phonemes or correct erroneous ones in real words by comparing results of bottom-up processes against lexical entries, or they can generate gross departures of perceptual experience from the stimulus as observed in mishearings.

The model in (12) accounts both for bottom-up analytic processes and a top-down constructions and restoration processes. They interact in the phonological working memory buffer where the structures underlying heard utterances are constructed.

Consider how new words are learned according to (12). First of all, in this case, the dictionary will not play any role in the analysis. The crucial component is the phonetic module that provides the featural inputs of the new words to the synthesis component that then constructs their complete representation. In partic-ular it hypothesizes possible underlying representation for the new word and then derives their surface representation by applying to it the processes (rules/repairs) of L1. In the case of foreign words with unfamiliar sounds, the featural input pro-vided by the phonological module cannot be used to construct licit featural repre-sentations of segments/syllabic configurations. The synthesis component must then adjust the featural input and construct representations that are licit according to the grammar of L1.

11. The construction of underlying representations

As proposed above, learning a word requires an analytical process that involves the grammatical knowledge that is used in production. In this section, I consider the analytical process involved in the construction of URs and demonstrate that the URs that are constructed in the case of foreign words must be consistent with L1 grammar. They must be "familiar", or "interpretable" in terms of grammar of L1,.

Before committing a foreign word to long-term memory, its underlying rep-resentation must be constructed. In Generative Grammar it is assumed that a UR ABC is postulated for a surface form AED in a language L when phonologi-cal alternations or distributional patterns in L provide evidence for two ordered processes in (13):
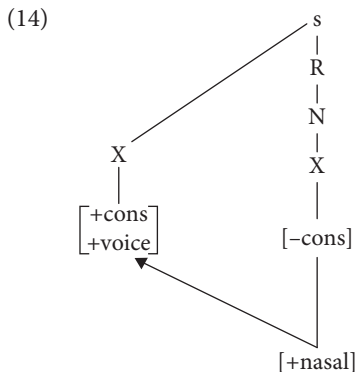
(13) a. $C \rightarrow D /$ ___ #
       b. $B \rightarrow E /$ ___ D

As is well known, postulating a base form implies postulating a rule, and viceversa, given the evidence provided by the alternations in the language. The L2 learner may have a limited access to the alternations needed to identify the UR of the L2 words. This limited access may be due to the time constraints of language acquisition or to the fact that as an adult the learner no longer has the ability to recognize and appropriately analyze the phonological alternations of L2, Therefore faced with a non-native form, the learner tends to analyze it in terms of his L1 system.

Given the analysis-by-synthesis model proposed here, the analysis of a word requires the reverse application of the phonological derivation. Therefore given the two rules in *(13)*, if the surface shape of the word is AED, an underlying ABC must be postulated so as to derive the surface AED.

In many cases the postulated UR does not need to be different from the surface L2 form. For example, if the L1 has an underlying distinction between voice vs. voiceless obstruents, and word-final devoicing, as in German, a speaker could assume an UR /gUd/ for English "good", despite the fact that he will pronounce this word [gUt].

In some cases, however, a different UR must be postulated. Take the Brazilian indigenous language Maxacalí. Wetzels (this volume) shows that Maxacalí nasality is contrastive only in the case of vowels. Nasal consonants are always derived by spreading the nasal feature of this vowel. In particular there is a rule spreading nasality from the vowel onto the syllabic onset of this vowel, making words such as *[bãŋõn] impossible in this language.

(14)

$$
\begin{array}{ccc}
 & & \text{s} \\
 & & | \\
 & & \text{R} \\
 & & | \\
 & & \text{N} \\
 & & | \\
\text{X} & & \text{X} \\
| & & | \\
\begin{bmatrix} +\text{cons} \\ +\text{voice} \end{bmatrix} & & [-\text{cons}] \\
 & [+\text{nasal}]
\end{array}
$$

Wetzels shows that in Brazilian Portuguese (BP) loanwords to Maxacalí, the original nasal onsets of the loanwords are analyzed as being the outcome of this spreading rule. As he puts it, "confronted with a nasal onset of an oral syllable, the speaker of Maxacalí interprets the nasal onset as an indication of the nasality of its nucleus." Therefore faced with BP words such as those in (15), a Maxacalí speaker

postulates a UR where the nasality is a property of the vowel. The rule in (14) then spreads the nasality onto the preceding voiced stop onset. The result is that the speaker postulates a UR consistent with the L1 phonological system.

(15)

		BP	Maxacalí
Margarida	'Margarida'	[mahgaɾida]	[mãⁿgadit]
carneiro	'sheep'	[kahnejɾω]	[kahnɛ̃n]
mesa	'table'	[mezɐ]	[mẽⁿdˀa]
moto	'motorbike'	[mɔtω]	[mõtok]

Observe that if the vowel is interpreted as [−nasal] in the borrowing, its onset is also non- nasal, just as expected if nasality is a property of the vowel and nasality in onset is derived by rule.

(16)

		BP	Maxacalí
martelo	'hammer'	[mahtɛlω]	[ᵐbahtɛt]
canivete	'pocketknife'	[kanivɛtˤi]	[kudibɛt]

Awareness of the rules and constraints of the L1 grammar, therefore, leads to the postulation of more abstract representations for L2, in particular the postulation of a representation for L2 consistent with the rules and constraints of L1. Consider some other examples. Nevins and Braun (this volume) discuss the following case involving the pronunciation of English by Brazilian Portuguese speaker in light of a BP rule changing the rhotic /r/ to a laryngeal fricative in word-initial position:

(17)   direto [dʒirɛtu]   vs.   reto [hetu]

Interestingly, in their pronunciation of English, Brazilian speakers replace word-initial /h/ with [r]:

(18)   BP pronunciation

home	[rom]	(or [hom])
hug	[rʌg]	(or [hʌg])
hunger	[rʌ̃gər]	(or [hʌ̃gər])

Nevins and Braun propose that when exposed to English words, a Brazilian learner observes that the rule debuccalizing [r] into [h] does not apply to English. When faced to word-initial /h/ in English, he then hypothesizes that it derives from underlying /r/ as in his own language. Given that he has postulated that r-debuccalization does not apply in English, this hypothesized /r/ surfaces in the English word as can be seen in (18). Again the UR postulated for these words is consistent with the grammar of L1.

    Another example involves my own pronunciation of English. Italian does not have the laryngeal fricative /h/. When I speak English, I delete it especially when word initial. However, I also often insert a laryngeal fricative in the same context.

(19)   My pronunciation
    harbor   [arbor]
    aisle   [hayl]

A possible analysis would be that in the UR of the relevant English words in my
long term memory, I do not have laryngeal fricatives as required by the Italian
grammar. Given that I observe that English has /h/, especially in word-initial
position, I hypothesize a rule of /h/-insertion to mask my problem. The point is
the UR of English h-initial words in my own lexicon is consistent with my own
native grammar.

12.   Galileo, Saturn and the pharyngeal vowels

In this concluding section, I will consider Jacobs and Gussenhoven's (2000) objec-
tion to the idea that a model that assumes that foreign sounds are modified in
perception predicts that people cannot hear segmental contrasts that do not occur
in their own language. According to this argument, if this is true, we would expect
major problems when the speaker of languages with small segment inventories,
like Tahitian or Maori are exposed to languages with larger inventories such as
English. For example, as the former languages lack the /t/ and /s/, one may pre-
dict that speakers of these languages would be incapable of hearing the difference
between these two segments given that their languages do not have this opposition.
However, as Jacobs and Gussenhoven observe, this is difficult to reconcile with the
common finding that language users appear to be capable of hearing (at least some)
non-native segments with ease. I agree with them. Kant also observed that raw
sensation is different from interpretation; i.e. having a sensation is different from
understanding it. A monolingual Tahitian or Maori speaker will hear that the
English /s/ and /t/ are different sounds. The issue is that he cannot understand how
they are produced, and therefore their identification remains fuzzy and unclear.

    With this in mind, I want to draw a parallel from the history of science to
explain this idea better. Eco (1977) conveys that when Galileo looked at Saturn for
the first time, he saw something never seen before. In his various letters to friends
and colleagues, Galileo described the efforts he made (as he looked) to understand
the shape of Saturn. For example, in three letters (to Benedetto Castelli, 1610, to
Belisario Giunti, 1610; and to Giuliano de Medici, 1611), he says he saw not one
star but three joined together in a straight line parallel to the equinoctial; he rep-
resented this in an drawing like the one below:

(20)   ◯ ◯ ◯

But in other letters (e.g. to Giuliano de Medici, 1610; and to Marco Velseri, 1612) he admits that owing "to the imperfection of the instrument and the eye of the observer", Saturn might also appear, as in (21), "in the shape of an olive".

(21)

The figure clearly reveals that, since it is wholly unexpected for a planet to be surrounded by a ring (which apart from anything else clashed with every notion held at the time with regards to heavenly bodies), Galileo was trying to understand what he could see; he was laboriously attempting to construct a (new) mental representation of Saturn.

After looking at the star and studying the situation for some time, (see his letter to Federigo Borromeo in 1610) Galileo finally decided that it was not a matter of two small round bodies but of larger bodies "and of a shape no longer round, but as can be seen in the enclosed figure, two semi-ellipses with two very obscure little triangles in the middle of the said figures, and contiguous to Saturn's middle globe." This consideration led Galileo to a third representation, (22):

(22)

Note that Galileo did not recognize the existence of rings, otherwise he would have written not of two semi-ellipses but of an elliptical band. It is only in trying to convey on paper the essential features of what he observed that Galileo gradually began to "see", to perceive Saturn and its rings. He finally "understood" its nature. Prior to that, Galileo could not recognize or identify what he was seeing, and he had to interpret it by trying different mental representations.

Observe that the sensation, or better the "sensory intuition", is still there before the interpretation. Our interpretation of the world does not change that. Therefore, Galileo was obviously able to distinguish what he was seeing in the case of Saturn from what he was seeing in the case of Jupiter. If, after his first viewing of Saturn, he had been asked if this new planet resembled Jupiter, he would have answered that it did not.

I claim that the same occurs in our perception and representation of foreign sounds. Once I attended a field method course with an Abhkaz speaker. In one of the classes, the informant uttered words with pharyngealized vowels. I had the distinct feeling that they were different from the plain ones, but I was not certain about the nature of this difference. Upon a second hearing, I mistakenly perceived them as fronted vowels, as diphthongs composed of a plain vowel plus something

else, and even as plain vowels. Only in later classes, after being told how to pronounce them, and having read some literature on the topic and having practiced pronouncing them, did I gain the ability to distinguish them from the plain ones in words that were pronounced slowly. Only after all this background could I begin to identify them, albeit only tentatively and in slow speech.

I wonder how to characterize this learning that had occurred to me. I heard these strange sounds, which were totally new to me (though admittedly there is a certain advantage in being a phonologist and knowing that pharyngealized vowels exist). While hearing utterances with these vowels, my phonological module constructed a first articulatory mental representation of these words, a sensory intuition. In the case of the pharyngealized vowels, the representations may have included the feature [+RTR]. In the phonological working memory buffer where phonology is accessed, incomplete representations or representation with the feature combination [−consonantal, +RTR] were illicit, and had to be adjusted by phonological operations, i.e. by repairs, in an attempt to produce a recognizable, familiar representation of these sounds, so as to apprehend them. These were illusory representations.

When, with training and effort, I learned to coordinate a [+RTR] configuration of the tongue root with a [−consonantal] stricture, I was able to match my internal representation of these sounds with their acoustic shape and I recognized them more or less well. But this was just tentative, and temporary, insofar as being an adult, I could not learn to articulate pharyngeal vowels, and I will always have both articulatory and perceptual problems with them. Obviously all of this was occurring in a very special context, a field method course, and I was being taught about pharyngeal vowels. Normal speakers are not so lucky and they will normally stop at the stage of the illusory representations. They may indeed feel that foreign sounds are auditorily different from other sounds, but they cannot identify or understand them because they cannot articulate them. In this case they will adjust them phonologically into sounds that are licit and articulatorily possible.

Articulating a sound and perceiving it, in the sense of apprehending it, is the same thing. To conclude with Giovann Battista Vico (1688–1744), *verum et factum reciprocantur seu convertuntur*. The human mind can know only what the human mind has made. Still, reality (acoustic reality, in the cases discussed here) is out there to check us, to control us, to stimulate us to change…

References

Aboitiz, Francisco & Ricardo V. Garcia. 1997. "The evolutionary origin of language areas in the human brain. A neuroanatomical perspective". *Brain Research Reviews* 25. 381–396.

Aslin, Richard N., Peter W. Jusczyk, & David B. Pisoni. 1998. "Speech and auditory processing during infancy: Constraints on and precursors to language". *Handbook of child psychology:*

Cognition, perception, and language ed. by Deanna Kuhn & Robert Siegler, 147–254. New York: Wiley.

Anderson, John. 1985. *Cognitive Psychology and Its Implications*. New York: Freeman.

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Awh Edward, John Jonides, Edward E. Smith, Eric H. Schumacher, Robert A. Koeppe & Stewart Katz. 1996. "Dissociation of storage and rehearsal in working memory: PET evidence". *Psychological Science* 7. 25–31.

Baddeley, Alan D. 1992. "Working memory". *Science* 255. 556–559.

Best, Catherine T. 1994. "The emergence of native language phonological influences in infants: A perceptual assimilation hypothesis". *The development of speech perception* ed. by Judith C. Goodman & Howard C. Nusbaum.167–224. Cambridge, MA: MIT Press.

Best, Catherine T. 1995. "A direct realist perspective on cross-language speech perception". *Crosslanguage speech perception* ed. by Winifred Strange & James J. Jenkins, 171–204. Timonium, MD: York Press.

Best, Catherine T., Gerald W. McRoberts, Rosemarie Lafleur, & Jean Silver-Isenstadt. 1995. "Divergent developmental patterns for infants' perception of two nonnative speech contrasts". *Infant Behavior and Development* 18. 339–350.

Blumstein Sheila & Kenneth N. Stevens. 1981. "Phonetic features and acoustic invariance in speech". *Cognition* 10. 25–32.

Boberg, Charles. 1999. "The Attitudinal Component of Variation in American English Foreign a". *Nativization, Journal of Language and Social Psychology* 18.1. 49–61.

Bobaljik, Jonathan D. 2006. "Itelmen Reduplication : Edge-in Association and Lexical Stratification". *Journal of Linguistics* 42. 1–23.

Boersma, Paul & Silke Hamann. This Volume. "Loanword adaptation as first-language phonological perception ".

Boring, Edwin G. 1946. "Perception of objects". *American Journal of Physics* 14. 99–107.

Bromberger, Sylvain & Morris Halle. 1992. "The ontology of Phonology". *What we know we don't know* by Sylvain Bromberger, 209–228. Chicago: Chicago University Press.

Bromberger, Sylvain & Morris Halle. 1997. "The content of phonological signs: a comparison between their use in Derivational Theories and in Optimality Theories". *Derivations and Constraints in Phonology* ed. by Iggy Roca, 93–122. Oxford: Oxford University Press.

Bromberger, Sylvain & Morris Halle. 2000. "The ontology of Phonology revised." *Phonological Knowledge. Conceptual and Empirical Issues* ed. by Noel Burton-Roberts, Philip Carr, & Gerald Docherty, 19–37. Oxford: Oxford University Press.

Browman, Catherine. 1980. "Perceptual processing: evidence from slips of the ear". *Errors in linguistic Performance: Slips of the tongue, ear, pen, and hand* ed. by Victoria Fromkin. New York: Academic Press.

Calabrese, Andrea. 1988. *Towards a theory of phonological alphabets*. Ph.D. diss., Massachusetts Institute of Technology.

Calabrese, Andrea. 1995. "A constraint-based theory of phonological markedness and simplification procedures". *Linguistic Inquiry* 26.2. 373–463.

Calabrese, Andrea. 2005. *Markedness and Economy in a Derivational Model of Phonology*, Mouton-De Gruyter, Berlin.

Chistovich, Ludmilla A. & Valentina V. Lublinskaya. 1979. "The 'center of gravity' effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel-like stimuli". *Hearing Research* 1. 185–195.

Dehaene-Lambertz, Ghislaine. 1997. "Electrophysiological correlates of categorical phoneme perception in adults". *NeuroReport* 84.3. 919–924

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux & A. Gout. 2000. "Electrophysiological correlates of phonological processing: a cross-linguistic study". *Journal of cognitive neuroscience* 12. 635–647.

Delgutte, Bertrand, & Nelson Y. S. Kiang. 1984. "Speech coding in the auditory nerve: IV. Sounds with consonant-like dynamic characteristics," *Journal of the Acoustic Society of America.* 75, 897–907.

Diehl, Randy L., Andrew J. Lotto, & Lori L. Holt. 2004. "Speech perception". *Annual Review of Psychology* 55. 149–179.

Di Pellegrino, Giuseppe, Luciano Fadiga, Leonardo Fogassi, Vittorio Gallese, & Giacomo Rizzolatti. 1992. "Understanding motor events: A neurophysiological study". *Experimental Brain Research* 91. 176–180.

Doupe, Allison J., & Patricia K. Kuhl. 1999. "Birdsong and human speech: common themes and mechanisms". *Annual Review of Neuroscience* 22. 567–631.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler. 1999. "Epenthetic vowels in Japanese: A perceptual illusion?". *Journal of Experimental Psychology: Human Perception and Performance* 25. 1568–1578.

Dupoux, Emmanuel, Christophe Pallier, Nuria Sebastian-Galles & Jacques Mehler. 1997. "A destressing "deafness" in French?". *Journal of Memory and Language* 36. 406–421.

Dupoux, Emmanuel, Christophe Pallier, Kazuhiko Kakehi & Jacques Mehler. 2001. "New evidence for prelexical phonological processing in word recognition". *Language and Cognitive Processes* 16. 491–505.

Dupoux, Emmanuel, Peperkamp, Sharon, & Nuria Sebastian-Galles (2001). "A robust method to study stress 'deafness'". *Journal of the Acoustical Society of America*, 110, 1606–1618.

Eco, Umberto. 1997. *Kant and the Platypus. Essays on Language and Cognition*. New York: Harcourt Brace.

Epstein, William. 1982. "Percept-percept couplings". *Perception* 111. 75–83

Fadiga, Luciano, Laila Craighero, Giovanni Buccino, & Giacomo Rizzolati. 2002. "Speech listening specifically modulates the excitability of tongue muscles: A TMS study". *European Journal of Neuroscience* 15. 399–402.

Fitzgerald, B.H. 1996. *Acoustic Analysis of Japanese Vowels Produced in Multiple Consonantal Contexts.* MA thesis, University of South Florida.

Flege, Jim E. 1991. "Perception and production: The relevance of phonetic input to L2 phonological learning." *Crosscurrents in second language acquisition and linguistic theories* ed. by Charles Ferguson & Thom Huebner. Philadelphia, PA: John Benjamins.

Fowler, Carol A. 1986. "An event approach to the study of speech perception from a direct-realist perspective". *Journal of Phonetics* 14. 3–28.

Fowler, Carol A. 1994. "Invariants, specifiers cues: An investigation of locus equations as information for place of articulation". *Perception* & *Psychophysics* 55. 597–610.

Fowler, Carol A. 1996. "Listeners do hear sounds, not tongues". *Journal of the Acoustical Society of America* 99. 1730–1741.

Fowler, Carol A. & Bruno Galantucci. 2005. "The relation of speech perception and production". *The handbook of speech perception* ed. by David B. Pisoni & Robert E. Remez. 633–652. Malden, MA: Blackwell.

Fowler, Carol A., Philip Rubin, Robert Remez, & Michael Turvey. 1980. "Implications for speech production of a general theory of action". *Language production: Vol.* 1. *Speech and talk* ed. by Brian Butterworth, 373–420. London: Academic Press.

Fox, Robert. 1984. "Effects of lexical status on phonetic categorization". *Journal of Experimental Psychology: Human Perception and Performance* 10. 526–540.

Galantucci, Bruno, Carol A. Fowler, & Michael T. Turvey. 2006. "The Motor Theory of Speech Perception Reviewed". *Psychonomic Bulletin and Review* 13.3. 361–377.

Ganong, William F. 1980. "Phonetic categorization in auditory word perception". *Journal of Experimental Psychology: Human Perception and Performance* 6. 110–125.

Garnes, Sara & Zinny Bond. 1980. "A slip of the ear: A snip of the ear? A slip of the year?". *Errors in Linguistic Performance: Slips of the tongue, ear, pen, and hand* ed. by Victoria Fromkin. New York: Academic Press.

Gibson, James J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gregory, Richard L. 1993. "Seeing & thinking". *Giornale Italiano di Psicologia* 205. 749–769

Grieser DiAnne & Patricia Kuhl. 1989. "Categorization of Speech by Infants: Support for Speech-Sound Prototypes". *Developmental Psychology* 25.4. 577–88

Goldstone Robert L. & Alan Kersten. 2003. "Concepts and Categorization". *Comprehensive Handbook of Psychology* ed. by Alice Healy & Robert Proctor, 599–621. New Jersey: Wiley

Goldinger, Stephen D. 1996. "Words and voices: Episodic traces in spoken word identification and recognition memory". *Journal of Experimental Psychology: Learning, Memory and Cognition* 22. 1166–1183.

Goldinger, Stephen D. 1998. "Echoes of echoes? An episodic theory of lexical access". *Psychological Review* 105. *251–279.*

Goldinger, Stephen D. 2000. "The role of perceptual episodes in lexical processing". *Proceedings of the Workshop on Spoken Word Access Processes SWAP* ed. by Ann Cutler, J.M. McQueen, & R. Zondervan, 155–159. Nijmegen: Max Plack Institute for Psycholinguistics.

Guion, Susan G. 2003. "The vowel systems of Quichua Spanish bilinguals: An investigation into age of acquisition effects on the mutual influence of the first and second languages". *Phonetica* 60. 98 -128.

Hale Mark & Charles Reiss. 1998. "Formal and Empirical Arguments concerning Phonological Acquisition". *Linguistic Inquiry* 29. 656–683 .

Halle, Morris. 1983. "On distinctive features and their articulatory implementation". *Natural Language and Linguistic Theory* 1.1. 91–105

Halle, Morris. 1995. "Feature geometry and feature spreading". *Linguistic Inquiry* 26. 1–46.

Halle, Morris. 2002. *From memory to speech and back. Papers on Phonetics and Phonology* 1954–2002. Berlin: Mouton De Gruyter.

Halle, Morris & Kenneth Stevens. 1962. "Speech recognition: A model and a program for research". *IEEE Transactions on Information Theory* 8. 155–159. Now in Halle 2002. 25–36.

Halle, Morris & Kenneth Stevens. 1991. "Knowledge of Language and the Sounds of Speech". *Music. Language, Speech and Brain* ed. by J. Sundberg et alii, 1–91. McMillan Press, London. Now in Halle 2002. 176–195.

Halle, Morris, Bert Vaux & Andrew Wolfe. 2000. "On feature spreading and the Representation of Place of Articulation". *Linguistic Inquiry* 31. 387–444.

Herd, Jonathan. 2005. "Loanword adaptation and the evaluation of similarity". *Toronto Working Papers in Linguistics* 24: *Special issue on similarity in phonology* ed. by Chiara Frigeni, Manami Hirayama & Sara Mackenzie, 65–116.

Helmholtz, Herman Von. 1867. Handbuch der Physiologischen Optik. Hamburg und Leipzig : Leopold Voss

Hickock, Gregory & David Poeppel. 2004. "Dorsal and Ventral Streams: a framework for understanding aspects of the functional anatomy of language". *Cognition* 92. 67–99.

Hintzman, Douglas 1986. "Schema abstraction in a multiple trace memory model". *Psychological Review* 93.411–428.

Homma, Yayoi 1981. "Durational relationship between Japanese stops and vowels". *Journal of Phonetics* 9. 273–281.

Houde, John F. & Michael I. Jordan. 1998. "Sensorimotor Adaptation in Speech Production." *Science* 20. 279. 1213–1216

Hyman, Larry 1970. "The role of borrowings in the justification of phonological grammars." *Studies in African linguistics* 1. 1–48.

Iacoboni, Marco, Roger P. Woods, Marcel Brass, Harold Bekkering, John C. Mazziotta & Giacomo Rizzolatti.1999. "Cortical mechanisms of human imitation". *Science* 286. 2526–2528.

Ingram, John & Park See-Gyoon. 1998. "Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners". *Journal of the Acoustical Society of America* 103. 1161–1174.

Ito, Chiyuki, Yoonjung Kang, & Michael Kenstowicz. 2006. "The Adaptation of Japanese Loanwords into Korean". *MIT Working Papers* 52. 65–104.

Iverson, Gregory & Ahrong Lee. 2004. "Perceived syllable structure in the adaptation of loanwords in Korean". Ms., University of Wisconsin, Milwaukee. [Presented at Workshop on Loanword Adaptation, 2004 Linguistic Society of Korea International Conference, Seoul, July 28–30]

Iverson, Paul, & Patricia K. Kuhl. 1995. "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling". *Journal of the Acoustical Society of America* 97. 553–562.

Iverson, Paul, & Patricia K. Kuhl. 2000. "Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism?". *Perception and Psychophysics* 62. 874–886.

Jacobs, Haike & Carlos Gussenhoven. 2000. "Loan phonology: perception, salience, the lexicon and OT". *Optimality Theory: Phonology, syntax, and acquisition* ed. by Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer, 193–210. Oxford: Oxford University Press.

Jakobson, Roman. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze* [Child Language, Aphasia and Phonological Universal]. Uppsala Universitets Aarskrift. Translated into English by A. Keiler. The Hague: Mouton, 1968.

Kabak, Baris & William Idsardi 2006. "Speech Perception is not isomorphic to phonology". ms. University of Konstanz, Konstanz.

Kanisza, Gaetano 1979. *The Organization in Vision: Essays on Gestalt Perception*, New York, Praeger.

Kenstowicz, Michael & Hyang-Sook Sohn. 2001. "Accentual adaptation in North Kyungsang Korean". *Ken Hale. A Life in Language* ed. by Michael Kenstowicz, 239–270. Cambridge, MA: MIT Press.

Kenstowicz , Michael. 1994. *Phonology in Generative Grammar*. Blackwell, Cambridge.

Kenstowicz, Michael. 2001. "The role of perception in loanword phonology". *Studies in African Linguistics* 32. 95–111.

Kenstowicz, Michael. 2003. "Salience and similarity in loanword adaptation: a case study from Fijian". *Language Sciences* 29. 316–340.

Kingston, John. 2003. "Learning Foreign Vowels language and speech". *Language and Speech* 46. 295–349.

Kim, Hyunsoon. This volume. "Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation".

Klatt, Dennis. 1980. "Speech perception: A model of acoustic-phonetic analysis and lexical access". *Perception and production of fluent speech* ed. by Ronald Cole. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Kohler Evelyne, Christian Keysers, M. Alessandra Umiltà, Leonardo Fogassi, Vittorio Gallese & Giacomo Rizzolatti. 2002. "Hearing sounds, understanding actions: Action representation in mirror neurons". *Science* 297. 846–848.

Kuhl, Patricia K. 1991. "Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not". *Perception and Psychophysics*, 50. 93–107.

Kuhl, Patricia K. 1993. "Innate predispositions and the effects of experience in speech perception: The native language magnet theory". *Developmental neurocognition: Speech and face processing in the first year of life* ed. by Benedicte de Boysson Bardies, Scania de Schonen, Peter Jusczyk, McNeilage, Peter & John Morton, 259–274. Dordrecht, Netherlands: Kluwer Academic Publishers.

Kuhl, Patricia K. 2000. "Language, mind, and brain: Experience alters perception." *The new cognitive neurosciences* ed. by Michael S. Gazzaniga, 99–115. Cambridge, MA: MIT Press.

Kuhl, Patricia K., Karen A. Williams, Francisco Lacerda, Kenneth N. Stevens, & Björn Lindblom. 1992. "Linguistic experience alters phonetic perception in infants by six months of age". *Science* 255. 606–608.

LaCharite, Darlene & Carole Paradis. 2005. "Category preservation and proximity vs. phonetic approximation in loanword adaptation". *Linguistic Inquiry* 36. 223–58.

Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages.* Oxford: Blackwell Publishers.

Liberman, Alvin M. 1957. "Some results of research on speech perception". *Journal of the Acoustical Society of America* 29. 117–123.

Liberman, Alvin M. 1996. *Speech: A special code.* Cambridge, MA: Bradford Books.

Liberman, Alvin M., Franklin S. Cooper, Donald Shankweiler & Michael Studdert-Kennedy. 1967. "Perception of the speech code". *Psychological Review* 74. 431–461.

Liberman, Alvin M., Pierre Delattre & Franklin S. Cooper. 1952. The role of selected stimulus variables in the perception of the unvoiced-stop consonants. *American Journal of Psychology,* 65. 497–516.

Liberman. Alvin M. & Ignatius G. Mattingly. 1985. "The motor theory or speech perception revised". *Cognition* 21. 1–36.

Liberman. Alvin M. & Ignatius G. Mattingly. 1989. "A specialization for speech perception". *Science* 243. 489–494.

Liberman, Alvin M., & Doug H. Whalen. 2000. "On the relation of speech to language". *Trends in Cognitive Sciences* 4. 187–196.

Mann, Virginia & Alvin Liberman. 1983. "Some differences between phonetic and auditory modes of perception". *Cognition* 14. 211–23.

Marr , David. 1982. *Vision.* San Francisco: Freeman.

Marslen-Wilson, William & Alan Welsh. 1978. "Processing interactions and lexical access during word recognition in continuous speech". *Cognitive Psychology* 10. 29–63.

McCawley, James 1968. *The Phonological Component of a Grammar of Japanese.* The Hague: Mouton.

McClelland, James & Jeff Elman. 1986. "The TRACE model of speech perception". *Cognitive Psychology* 18.1–86.

Merleau-Ponty, Maurice. 1945. *Phénoménologie de la perception*, Paris: Gallimard, 1945.Trans. by Colin Smith, (New York: Humanities Press, 1962.

Milner, A. David & Melvyn A. Goodale 1995. *The visual brain in action*. Oxford: Oxford University Press.

Näätänen, Risto. 2001. "The perception of speech sounds by the human brain as reflected by the mismatch negativity MMN and its magnetic equivalent MMNm". *Psychophysiology* 38. 1–21.

Neisser, Ulric. 1967. *Cognitive psychology*. Appleton-Century-Crofts New York.

Nevins, Andrew & David Braun. This Volume. "The role of underlying representations in L2 Brazilian English".

Norman, Joel. 2001. "Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches". *Behavioral and Brain Sciences* 25. 73–96.

Okada, Hideo 1999. "Japanese". *Handbook of the International Phonetic Association:* A *guide to the usage of the International Phonetic Alphabet.* 117–119 Cambridge: Cambridge University Press.

Paavilainen, Petri, M. Jaramillo, Risto Näätänen, & Istvan Winkler. 1999. "Neuronal populations in the human brain extracting invariant relationships from acoustic variance". *Neuroscience Letters* 265. 179–182.

Paavilainen, Petri, J. Simola, M. Jaramillo, Risto Näätänen & Istvan Winkler. 2001. "Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity MMN". *Psychophysiology* 38. 359–365.

Pallier, Christophe, Laura Bosch, Nuria Sebastian Gallés. 1997. "A limit on behavioral plasticity in speech perception". *Cognition* 643. 9–17.

Paradis, Carole. 1992. *Lexical Phonology and Morphology: The Nominal Classes in A Psycholinguistic Theory of Loanword Adaptations in Fula.* New York: Garland.

Paradis, Carole. 1995. "Native and loanword phonology as one: constraints versus rules". *Proceedings of the 13th International Congress of Phonetic Sciences*, ed. by K. Elenius & P. Branderud, 374–81. Stockholm.

Paradis, Carole & Antoine Tremblay. This Volume. "Nondistinctive Features in Loanword Adaptation: The Unimportance of English Aspiration in Mandarin Chinese Phoneme Categorization."

Peperkamp, Sharon. 2002. "Towards a new theory of loanword adaptations". Ms., University of Paris 8.

Peperkamp, Sharon, & Emmanuel Dupoux. 2002. "A typological study of stress "deafness"." *Laboratory phonology 7* ed.by Carlos Gussenhoven & Natasha Warner, 203–240. Berlin: Mouton de Gruyter.

Peperkamp, Sharon & Emmanuel Dupoux. 2003. "Reinterpreting loanword adaptations: the role of perception". *Proceedings of the I5th International Congress of Phonetic Sciences.* 367–70.

Phillips, Colin. 2001. "Levels of representation in the electrophysiology of speech perception". *Cognitive Science 25.* 711–731.

Phillips, Colin, Thomas Pellathy, Alec Marantz, Elron Yellin, Ken Wexler, David Poeppel, Martha McGinnis, Timothy Roberts. 2000. "Auditory Cortex Accesses Phonological Categories: An MEG Mismatch Study". *Journal of Cognitive Neuroscience* 12. 1038–1055.

Piaget Jean. 1937. *La construction du reel chez l'enfant*. Neuchatel: Delachaux et Niestle

Piaget, Jean. 1969. *Mechanisms of perception* New York: Basic Books.

Polka, Linda, & Janet F. Werker. 1994. "Developmental changes in perception of nonnative vowel contrasts". *Journal of Experimental Psychology, Human Perception, and Performance* 20. 421–435.

Pulvermüller, Friedemann, Martina Huss, Ferath Kherif, Fermin Moscoso Del Prado Martin, Olaf Hauk & Yuri Shtyrov. 2006. "Motor cortex maps articulatory features of speech sounds". *Proceedings of the National Academy of Sciences* 103. 7865–7870.

Pulvermüller, Friedemann, Teja Kujala, Yuri Shtyrov, Jaana Simola, Hannu Tiitinen, Paavo Alku, Kimmo Alho, S. Martinkauppi, Risto J. Ilmoniemi, & Risto Näätänen. 2001. "Memory traces for words as revealed by the Mismatch Negativity MMN". *NeuroImage* 14. 107–616.

Repetti, Lori. 2003. "Come i sostantivi inglesi diventano italiani: la morfologia e la fonologia dei prestiti". *Italiano ed Inglese a confronto*, ed. by Sullam Calimani, 31–42. Firenze: Giunti.

Repetti, Lori. 2006. "The emergence of Marked Structures in the Integration of Loans in Italian" *Historical Romance Linguistics. Retrospectives and Perspectives*, ed. by Randall S. Gess & Deborah Arteaga. Amsterdam John Benjamins.

Repetti, Lori. This volume. "Gemination in English loans in American varieties of Italian"

Rizzolatti, Giacomo & Michael A. Arbib. 1998. "Language within our grasp". *Trends in Neurosciences* 21. 188–194.

Rizzolatti, Giacomo, R. Camarda, Leonardo Fogassi, M. Gentilucci, G. Luppino, & Massimo Matelli. 1988. "Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements". *Experimental Brain Research* 71. 491–507.

Rizzolatti, Giacomo, & Laila Craighero. 2004. "The mirror-neuron system". *Annual Review of Neuroscience* 27. 169–192.

Rizzolatti, Giacomo, Leonardo Fogassi, & Vittorio Gallese. 2001. "Neurophysiological mechanisms underlying the understanding and imitation of action". *Nature Reviews Neuroscience* 2. 661–670.

Rock, Irving. 1977. "In defense of unconscious inference." *Stability and constancy in visual perception: Mechanisms and processes* ed. by William Epstein, 321–377. New York.

Rock, Irving. 1983. *The logic of perception*. Cambridge, MA: MIT Press.

Rock, Irving. 1997. *Indirect perception*. Cambridge, MA: MIT Press.

Samuel Arthur G. 1981. "Phonemic Restoration: Insights from a New Methodology." *Journal of Experimental Psychology* 110. 474–494.

Sapir, Edward. 1933. "The Psychological Reality of Phonemes". *Selected Writings of Edward Sapir in Language, Culture and Personality* ed. by David G. Mandelbaum, 46–60. Berkeley: University of California Press. 1949.

Shestakova, Anna; Elvira Brattico, Minna Huotilainen, Valery Galunov, Alexei Soloviev, Mikko Sams, Risto J. Ilmoniemi, Risto Näätänen. 2002. "Abstract phoneme representations in the left temporal cortex: magnetic mismatch negativity study". *Neuroreport* 1314. 1813–1816.

Shinohara, Shikego. 1997. *Analyse phonologique de l'adaptation japonaise de mots etrangers*. Doctoral dissertation, Universite de Paris III.

Shinohara, Shikego. 2000. "Default accentuation and foot structure in Japanese: evidence from Japanese adaptations of French words". *Journal of East Asian Linguistics* 9. 55–96.

Silverman, Daniel. 1992. "Multiple scansions in loanword phonology: evidence from Cantonese". *Phonology* 9. 289–328.

Smith, Jennifer. 2005. "Loan Phonology is not All Perception: Evidence from Japanese Loan Doublets". *Japanese/Korean Linguistics 14* ed. by Timothy J. Vance. Stanford: CLSI.

Stevens, Kenneth N. 1972. "The quantal nature of speech: Evidence from articulatory-acoustic data". *Human Communication: A Unified View* ed. by Eduard E. David & Peter B. Denes, 51–66. New York: McGraw-Hill.

Stevens, Kenneth N. 1989. "On the quantal nature of speech," *Journal of Phonetics* 17. 3–46.

Stevens, Kenneth N. 1998. *Acoustic Phonetics.* Cambridge, MA: MIT Press.

Stevens, Kenneth N. 2002. "Toward a model for lexical access based on acoustic landmarks and distinctive features". *Journal of the Acoustical Society of America* 111 4. 1872–1891.

Stevens, Kenneth N. & Blumstein, Sheila. 1978. Invariant cues for place of articulation in stop consonants, *Journal of the Acoustical Society of America,* 64, 1358–1368.

Stevens, Kenneth N. & Blumstein, Sheila. 1981. "The search for invariant correlates of phonetic features." *Perspectives on the study of speech* ed. by Peter D. Eimas & Joanne L. Miller. Hillsdale, NJ: Erlbaum.

Sussman, Harvey, Fruchter, D., Hilbert, J., & Sirosh, J. 1999. "Linear correlates in the speech signal: The orderly output constraint." *Behavioral and Brain Sciences* 21. 287–299.

Tervaniemi, Mari, J. Saarinen, Petri Paavilainen, N. Danilova, & Risto Näätänen. 1994. "Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity". *Biological Psychology* 38. 157–167.

Ungerleider, Leslie G. & Mortimer Mishkin. 1982. "Two cortical visual systems". *Analysis of Visual Behavior* ed. by David J. Ingle, Melvyn A. Goodale & Richard J.W. Mansfield, 549–586. Cambridge, MA.

Valdman, Albert. 1973. "Some Aspects of decreolization in Creole French". *Diachronic, Areal, and Typological Linguistics. Current Trends in Linguistics, vol. II* ed. by Thomas Sebeok. The Hague: Mouton.

Van Leuween, Cees. 1998. "Perception". *A Companion to Cognitive Sciences*. ed. by William Bechtel & George Graham, 265–282. Oxford: Blackwell Publishing.

Vance, Timothy J. 1987. *An Introduction to Japanese Phonology.* Albany, NY: State University of New York Press.

Waldstein, Robin S. 1989. "Effects of postlingual deafness on speech production: implications for the role of auditory feedback". *Journal of the Acoustical Society of America* 88. 2099–2144.

Warren, Richard. 1970. "Perceptual restoration of missing speech sounds". *Science* 167. 392–393.

Watkins, K. E., Strafella, A. P., & Paus, T. 2003. "Seeing and hearing speech excites the motor system involved in speech production". *Neuropsychologia* 41. 989–994.

Werker, Janet F., & Chris E. Lalonde. 1988. "Crosslanguage speech perception: Initial capabilities and developmental change". *Developmental Psychology* 24. 672–83

Werker, Janet F., & Richard C. Tees. 1984. "Crosslanguage speech perception: Evidence for perceptual reorganization". *Infant Behavior and Development* 7. 49–63.

Wetzels, Leo. This volume. "Nasal harmony and the representation of nasality in Maxacalí: evidence from Portuguese loans".

Wilson, Margaret. 2001. "The case for sensorimotor coding in working memory". *Psychonomic Bulletin and Review* 8.44–57.

Wilson, Stephen M., Ayse P. Saygin, Martin I. Sereno, & Marco Iacoboni. 2004. "Listening to speech activates motor areas involved in speech production". *Nature Neuroscience*. 7. 701–702.

Woldorff, Marty G., Steven A Hackley & Stephen A Hillyard. 1991. "The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones".*Psychophysiology* 28.1. 30–42.

Yates, Aubrey J. 1963. "Delayed auditory feedback". *Psychological Bulletin* 60. 213–251.

Yip, Moira. 1993. "Cantonese loanword phonology and Optimality Theory". *Journal of East Asian Linguistics* 2. 261–291.

The adaptation of Romanian loanwords from Turkish and French*

Michael L. Friesner
University of Pennsylvania

This paper examines several factors affecting loanword adaptation, using a data set of Romanian loanwords from Turkish and French. After exploring the position of loanwords in the lexicon and the nature of the two contact situations, the author considers relevant social, morphological, and phonological factors. First is the difference in the loanwords' semantic domains and their motivations for being borrowed. Next, the author introduces the morphophonological factors considered—stress, desinence class, and gender assignment—and discusses their behavior in the core vocabulary and previous relevant studies. Subsequently, the author examines the loanword data in detail, comparing and contrasting the Turkish- and French-origin loanwords. The author concludes that one must consider different modules of the language—the phonology and the morphology—and that only by contrasting borrowings from different languages into the same language can one determine the relative effect of internal and external factors on the outcome of contact.

1.   Introduction

The issue of the nativization of loanwords has been discussed in terms of a 'core-periphery' organization of the lexicon (cf. Itô & Mester 1995a,b). Such a model suggests that peripheral lexical items may be exceptional with regard to certain constraints of the recipient language. The typical path for a foreign borrowing is thus to enter the language in the periphery and then optionally to become fully or partially nativized, usually by changing its surface form to obey the previously violated constraints.

Loanword adaptation is frequently studied in terms of the phonology alone. In this paper, I consider themes that examine more broadly the question of how

loanwords are nativized. What are the internal and external factors that influence loanword adaptation patterns, and how do we assess the relative importance of these factors? How does the interaction between morphology and phonology come into play in loanword adaptation?

In considering the particular case of Romanian, I examine the effects of contact with French and Turkish, drawing from a collection of commonly used loanwords from these two languages, which I compiled. In particular, I address to what extent the differences in adaptation patterns of loanwords from Turkish and those from French can be explained by external, as opposed to internal, factors. The nature and phonological shape of these borrowings can thus be explained in large part by phonological and morphological considerations, but these must be coupled with an examination of the type of contact involved.

1.1   Loanwords and the lexicon

As mentioned earlier, Itô & Mester's (1995a,b) model of a 'core-periphery' organization of the lexicon suggests that peripheral items are allowed to violate constraints that are active in the core. Peripheral items include proper names, specialized vocabulary, onomatopoetic forms, and, most notably, words of foreign origin. The typical path for a foreign borrowing is thus to enter the language in the periphery. These borrowings may then eventually become part of the core, thus coming to obey all the constraints of the language, or they may remain in the periphery, despite being partially nativized. A fully nativized core lexical item is not perceived as foreign or exceptional by native speakers.

In Optimality Theory, this change in surface form is accounted for by the reranking of Faithfulness constraints (cf. Davidson & Noyer 1997). For peripheral items, if the relevant Faithfulness constraints are ranked above the relevant Markedness constraints, then the result is more faithful to the input from the source language.

Some languages presumably have a more distinctly stratified lexicon than others. The prime example used by Itô & Mester (1995a,b) is that of Japanese, which, due to its history, has distinct strata of the lexicon that correspond to native Japanese vocabulary, early Chinese loans ('Sino-Japanese'), and more recent loans that can vary in degree of nativization. These authors demonstrate that the strata can be distinguished by specific constraints, which all apply to the core vocabulary. In subsequent peripheral layers, the constraints are increasingly allowed to be violated, as they are ranked lower than Faithfulness constraints. This model is demonstrated in Figure 1.

The Romanian lexicon can be described as having a similar core-periphery structure to that of Japanese. While Romanian developed from Latin, its core

vocabulary contains a large percentage of words of Slavic origin, as well as other early borrowings from Turkish, Hungarian, German, Greek, and Albanian, among other languages (Chitoran 2001). These were the earliest languages with which Romanian was in contact. In particular, the Slavic contact yielded borrowings from the fifth century CE onwards (Petrucci 1999). Close (1974), claims that this early contact resulted in the borrowing of lexical items that had been fully assimilated by the sixteenth century. Maneca (1966) shows that borrowing accounts for the finding that the Latin core constitutes only about 35% of the lexicon of modern Romanian.
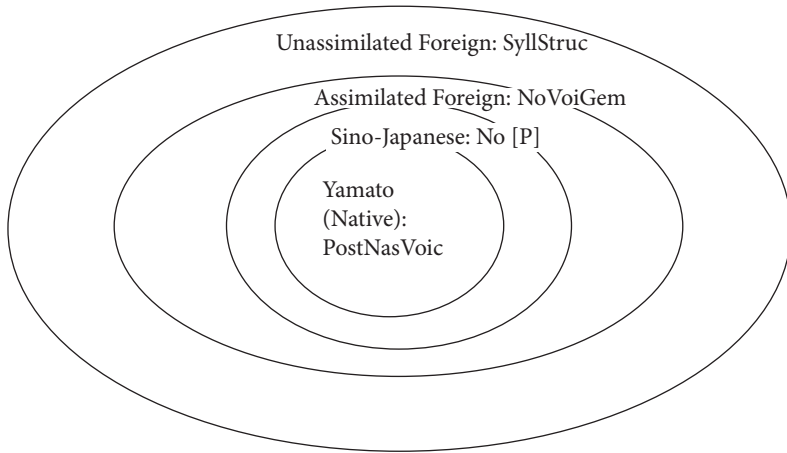

Figure 1. Itô & Mester's (1995b) core-periphery model of the Japanese lexicon

Based on the distinct waves of contact described above, Chitoran (2001:31–32) suggests the structure of the Romanian lexicon shown in (1).

(1) *native core vocabulary*:   Latinate vocabulary
     other core vocabulary: Slavic and other early loans
     partly-assimilated vocabulary: French, Italian, Greek, and Turkish loans
                                       from the 14th to 19th centuries
     unassimilated vocabulary: recent English loans

This paper focuses on Chitoran's (2001) category of partially assimilated loans and suggests that this category may need to be further subdivided.

1.2   The nature of the two contact situations

Turkish contact with Romanian was at its height between the fourteenth and eighteenth centuries (Close 1974; Chitoran 2001). This contact was due primarily to the expansion of the geographic area of Turkish control. Many of the loanwords

from Turkish from this period have fallen out of use in modern Romanian, a fact which was facilitated by conscious efforts of Romanian intellectuals to eliminate Turkish 'impurities' from the Romanian lexicon. This amazingly successful corpus planning effort constituted an assertion of independence. The Turkish words examined here are among the few that have remained in use subsequent to this 'purification' process.

Contact with French was at its height during the nineteenth century. This contact was of a very different nature. French was the language of intellectualism and sophistication. Borrowing from French, as well as Italian, served as a way to adopt Latinate lexical items to replace earlier Turkish borrowings. This was a deliberate way to enrich the vocabulary of Romanian with words consistent with its Romance roots. Some French loanwords were adopted and adapted from writing, in a conscious manner, and were initially used primarily within the upper tiers of society.

1.3   The data set

The data considered here come from a compilation of Romanian loanwords from Turkish and French. Only those words which are still in use today were included from a much larger set.[1] The exclusions were intended to allow verification of actual pronunciation, as prescriptive pronunciations of loanwords often differ from their most frequent pronunciation by native speakers. Words of unclear origin were also excluded from the data set,[2] as were clear neologisms and learned forms, as much as possible. Other words which have been excluded include so-called 'international loans', which may have several sources, all of which may have interacted to produce the Romanian surface form (Close 1974:38–39). The analysis here is thus based on the eighty-five relatively frequently occurring forms from the data set that remain after these exclusions.

An additional concern is the possible effect of orthography. While this is a legitimate concern, the phonological features examined here have been selected in part to minimize the possibility of such effects. These features—stress, desinence class assignment, and gender assignment—are less likely to exhibit orthographic effects than segmental features.

---

1.   Romanian data are drawn from Chitoran (2001), Close (1974), Sala (1976), and Suciu (1992). Only borrowings cited in the source material which were familiar to two native speaker informants were included for analysis.

2.   It is often difficult to distinguish words that came into Romanian from French and those that were borrowed from Italian. I thus excluded words where adaptation of the French and Italian form would likely have yielded the same result.

2.   Semantic domains of loanwords

The differences in nature of contact are manifested in the semantic domains of the loanwords from the data set and the type of semantic shift they undergo. Many authors discuss the various motivations for borrowing (Weinreich 1968; Haugen 1969; Poplack, Sankoff & Miller 1988; *inter alia*). These include the need to fill a lexical gap, the desire to adopt a prestige form, the desire for a more localized term or one that carries covert prestige, and the usefulness of a more succint or morphologically simpler way to express a concept. However, depending on the contact situation, the relative importance of each of these possible factors may differ. Attitudes towards the source language and its speakers, influenced by factors such as the nature of the contact situation and cultural prejudices, as well as which members of the community are likely to adopt the borrowings first, affect the type of semantic shift that tends to occur with loanwords.

Loanwords in the data set usually fall within the expected semantic domains for loanwords: food, drink, household items, and materials. Some examples of such loanwords from the data set are given in (2).

(2)	a.	Turk.	tʃórba 'soup'	> Rom.	tʃórbə
	b.	Turk.	kahvé 'coffee'	> Rom.	kaféa
	c.	Fren.	pyré 'purée'	> Rom.	pjuré/piréw
	d.	Turk.	kanepé 'couch'	> Rom.	kanapeá
	e.	Turk.	basmá 'cloth'	> Rom.	basmá 'scarf'
	f.	Turk.	perdé 'curtain'	> Rom.	perdéa
	g.	Fren.	vwál 'veil'	> Rom.	vwál
	h.	Fren.	ʒaluzí 'Venetian blind'	> Rom.	ʒaluzéa

Even within this realm, words of Turkish origin *tend* to refer to more commonplace objects, while the French words have more specific, high-end uses.

Other semantic domains reflect more clearly the differences in nature of contact. Turkish loans tend to refer to aspects of the government and the military, as shown in (3). This is not unexpected given that for a time these bodies were controlled by the Turks. The sources show that there were many more words within these domains that are no longer used.

| (3) | a. | Turk. | aɣá 'government official' | > Rom. | áɡə |
| | b. | Turk. | paʃá 'general' | > Rom. | páʃə |

French loans, on the other hand, tend to refer to aspects of high society, as in (4).

(4)	a.	Fren.	budwár 'boudoir'	> Rom.	budwár
	b.	Fren.	lɔrɲét 'opera glasses'	> Rom.	lornétə
	c.	Fren.	barɔ̃́ 'baron'	> Rom.	barón

While many of the words were adopted out of necessity, to fill a semantic gap, others exist alongside a native word. In such instances, we expect some kind of differentiation between the two lexical items according to register or connotation. In fact, the French loans tend to be attributed a positive connotation, while the Turkish loans take on a negative connotation. Some examples are given in (5).

(5) a. Fren. dam 'lady'   > Rom. dámə
     b. Fren. bal 'dance/ball'   > Rom. bál
     c. Fren. balkɔ́ 'balcony'   > Rom. balkón
     d. Turk. jaɣmá 'loot/pillaging' > Rom. jámə
     e. Turk. belá 'trouble'   > Rom. belʲá
     f. Turk. kelepír 'bad bargain' > Rom. kilipír

In the most extreme cases, a semantic shift occurs in the direction of linguistic attitudes. These words' trajectories reflect speakers' attitudes toward the source language and culture. The most striking examples involve the pejorative meaning shift of several of the Turkish loanwords, such as those shown in (6).

(6) a. Turk. hajmaná 'wandering'   > Rom. hajmaná 'tramp'
     b. Turk. mahallé 'neighborhood' > Rom. mahalá 'slum'

3. Stress and gender desinence

3.1 Stress and gender desinence in the native vocabulary

In the native vocabulary of Romanian, stress can surface anywhere from the final syllable to the preantepenultimate. As Petrucci (1999:39–41) explains, the quantity sensitive system that existed in Latin was gradually lost in Romanian due to processes of vowel shortening and loss of unstressed syllables. Looking at the internal structure of words in modern Romanian, stress assignment can be said to follow two patterns: stem-penultimate or stem-final (cf. Chitoran 2001; Friesner 2006). Under this analysis, gender desinences and certain suffixes are excluded from the domain of stress, thus yielding the other surface stress patterns. Some examples are given in (7).

(7) a. [kə.már]ə 'pantry'   b. [mə.se̯á] 'tooth'   c. [ma.gérn]its-ə 'hovel'
     d. [ká.mer]ə 'room'   e. [rúp-e] 'he tears' f. [vé.ver]its-ə 'squirrel'

Desinence vowels generally correspond to a specific gender. Following Harris (1991) and Repetti (2003, 2006), I treat these desinences as declension classes that are assigned a certain gender.[3] Chitoran (2001) explains that nearly all native nouns in Romanian end in one of the desinence vowels.

---

3. I sometimes refer to these as 'gender desinences' for the sake of simplicity.

As shown in (8), feminine nouns generally end in -ə or -e.

(8)   a.   kás-ə 'house'   b.   kárt-e 'book'

Masculine/neuter nouns[4] usually end in -e or -u underlyingly. Underlying -u is usually deleted, except when the result would be ill-formed, but its presence in consonant-final masculine nouns is convincingly justified by Chitoran (2001:37–39) and Iscrulescu (2003). This final -u may also sometimes be realized as -w, after a vowel. Some examples are given in (9).

(9)   a.   múnt-e 'mountain'   b.   bívol(-u) 'buffalo'
      c.   kúpl-u 'couple'   d.   karó-w 'square'

Almost all native vocabulary bears a desinence vowel, but there are about a dozen native words descended from Latin which lack a desinence vowel (10).

(10)   a.   steá 'star'   b.   məseá 'tooth'   c.   purtʃeá 'female pig'

Given the rarity of words lacking a gender desinence vowel, the examples in (10) seem to reflect a marked pattern. However, such words may serve as a basis for analogy in loanword adaptation.

Petrucci (1999) shows that early Slavic loans were assigned a desinence class, but they retained the stress position from the source with occasional concomitant consonant deletion. The only exception to this pattern was final-stress words, which underwent a shift in stress, as shown in (11).

(11)   a.   Slavic slugá 'servant' > Rom. slúgə
       b.   Slavic xraná 'food'   > Rom. hránə

This shift allowed the Slavic-origin words to fall into one of the native patterns of stem-final or stem-penultimate stress.

3.2   Studies of stress and gender desinence in loanwords

Stress assignment in loanwords has been studied in a number of languages, including English (Svensson 2001), Huave (Davidson & Noyer 1997), Kyungsang Korean (Kentowicz & Sohn 2001), Thai (Kenstowicz & Suchato 2006), and Fijian (Kenstowicz 2007), as well as cross-linguistically by Peperkamp & Dupoux (2002). These studies generally suggest two possible outcomes for loanword stress assignment: maintenance of stress position from the source language or adaptation to the unmarked stress position of the recipient language. In terms of a core-periphery model, these

---

4.   'Neuter' nouns behave like masculine nouns in the singular and feminine nouns in the plural. The status of 'neuter' as a separate gender in Romanian is under debate. Here, where only singular forms are considered, these nouns are simply referred to as 'masculine'.

two possibilites could be restated as non-adaptation and full adaptation. These outcomes can be accounted for within Optimality Theory through the reranking of FAITH(stress) relative to the relevant markedness constraint for stress in the recipient language.

An impressionistic look at the data from a number of language pairs paints a more complex picture. For example, Santorini (p.c.) notes that some trisyllabic loanwords from French into Middle English surface with penultimate stress. This constitutes neither non-adaptation, which would yield final stress, nor full adaptation, which would yield initial stress. Thus, there must be some other possible intervening factors. These may include the presence of secondary stress in the source form, analogy with other lexical items in the native vocabulary of the recipient language, or the need for loanwords to adhere to morphological requirements of the recipient language.

Gender assignment in loanwords has not been examined as extensively as stress. The existing studies addressing this issue (e.g., Fisiak 1975; Poplack, Pousada & Sankoff 1982; Thornton 2003) indicate a complex interplay of factors, including semantic effects, various types of analogy, and orthographic influences. The interaction between such morphological factors and phonological factors in loanword adaptation is rarely considered (but note Repetti 2003, 2006, this volume). Questions related to such interactions are relevant to the analysis of the data presented here.

4.   Stress and desinence vowels in the Turkish loanwords

4.1   Gender desinence in the Turkish loans

In considering the gender assignment of Turkish loanwords, it is important to note that nouns in Turkish do not carry gender. Thus, Turkish loans in Romanian are necessarily assigned to a declension class without the influence of the grammatical gender from the source language. As a result, desinence class assignment is made based primarily on phonological form. Interestingly, natural gender does not seem to play a role in desinence class assignment (cf. páʃə 'general'), although it does affect gender assignment as manifest in adjective agreement, for example. There are some examples of masculine nouns with a feminine desinence vowel in the native vocabulary, as well.

As shown in (12), consonant-final nouns are assigned masculine gender and are treated as if they have undergone the usual -*u* deletion that occurs with masculine nouns.

(12) a.   Turk.   kelepír 'bad bargain'   >   Rom.   kilipír
      b.   Turk.   gavúr 'foreigner/infidel' >   Rom.   gjaúr

Vowel-final nouns in the data set are assigned to a feminine gender declension in Romanian, as shown in (13). All such examples in the data set end in /a/ or /e/ in Turkish, except for one that ends in /y/. More specifically, Turkish final /a/ is generally adapted either as stressed /á/ or else as /ə/ with concomitant stress shift. Turkish /e/, on the other hand, is generally adapted as /ęá/. Turkish /e/-final words are not assigned to the -*e* feminine gender desinence. This seems to constitute a closed declension class.

(13) a.   Turk.   sobá 'stove'   >   Rom.   sóbə
      b.   Turk.   basmá 'cloth'   >   Rom.   basmá 'scarf'
      c.   Turk.   kadifé 'velvet'   >   Rom.   katifęá
      d.   Turk.   gøtyrý 'price/lump'   >   Rom.   gjóturə 'a lot'

The difference observed in the two adaptation patterns for Turkish final stressed /a/ seems to reflect a tendency for /a/ to be adapted as the /ə/ desinence in older borrowings, and for /a/ to be adapted as final stressed /a/ in newer borrowings. This latter pattern is closer to the Turkish pronunciation, but it constitutes a category of words not attested in the native vocabulary of Romanian.

This patterning recalls Haugen's (1950) analysis of American Norwegian in contact with English. Haugen found that later borrowing generations allowed loanwords to remain less adapted because there was greater familiarity with input from the source language that followed this nonnative pattern. Thus, the pattern seemed less exceptional to the native speakers of the recipient language. However, some counterevidence to this explanation comes from the existence of words which have recently changed in form while being nativized to a greater degree, such as the example given in (14).

(14) Turk.   paʃá 'general' >   Rom.   paʃá >   newer Rom.   páʃə

While there are very few exceptions to the pattern of /a/ being adapted as /ə/ or /a/ and /e/ as /ęa/, some do exist, as given in (15).

(15) a.   Turk.   belá 'trouble' >   Rom.   belęá
      b.   Turk.   kulé 'tower'   >   Rom.   kúlə

These exceptions (only three in the data set) could be the result of imperfect learning, where the Turkish form is misconstrued. Alternatively, these frequent words may in some cases be perceived as more native, and thus they are assigned to a more native-like pattern than would be expected. Final stressed /ęá/, for example, has a basis in the native vocabulary, while /á/ does not.

4.2   Stress in the Turkish loans

In the native vocabulary, Turkish stress almost always falls on the final syllable. Two exceptions to this tendency are place names and certain lexical items, which are generally described in the literature to be lexically marked (Underhill 1976:18–19). Other non-final surface stress is explainable in terms of Turkish morphophonology.

The stress position from Turkish is not always maintained in loanword adaptation. This outcome seems to suggest that the need for an overt gender desinence marker can override faithfulness to input stress. In Optimality Theory, this finding can be captured through the ranking of a constraint such as REALIZE-MORPHEME(gender) (cf. Walker 2000) above FAITH(stress).[5]

However, in the cases where stress from Turkish is maintained without the addition of a desinence vowel, it may be that the existence of a few native words with final stress that lack a desinence vowel (ending in /ęa/) serves as a basis of analogy. If this is so, this would explain the slight preference for adaptations in /ęa/, an ending attested in the native vocabulary.

Nonetheless, I do not argue that these words should be analyzed *identically* to the exceptional final-stress core vocabulary items. The borrowings still seem peripheral (cf. Itô & Mester 1995a,b) in a way that the native words may not, in that they are still often known to be borrowed items. If they *were* joining the exceptional native category, the prediction would be that they would remain exceptions even when being nativized to the point of being imperceptibly foreign in origin. Instead, the examples of such words that have undergone further nativization indicate that they change in form to adhere to the dominant core pattern, with a desinence vowel (as in (14)). The analysis suggested here for the final-stress words in Romanian is thus similar to that I proposed in Friesner (2001) for 'h aspiré' words in French. These are an exceptional class of consonant-like vowel-initial words, usually with an initial *h* in the orthography. This is the category to which most recent *h*-initial loans are assigned, but there is a basis in the core vocabulary, descending from either Latin *f*-initial words or early Germanic loans beginning with *h*. In Friesner (2001), I found that when these words were perceived as native (e.g., for many speakers, *handicapé* 'handicapped', *hovercraft*, and *hamburger*), they were likely to be treated as unexceptional, vowel-initial words. Native exceptional words, on the other hand, are somehow lexically marked to allow the exceptional

---

5.   As one reviewer points out, other constraints must rank even higher to account for the deletion of -*u* that occurs in native masculine forms ending in a consonant. As a full Optimality Theory analysis is beyond the scope of this paper, I will leave consideration of the exact formalism needed to account for these data to future work.

behavior, thus constituting an island that otherwise adheres to the core constraints of the language. The model to account for this behavior is given in Figure 2, from Friesner (2001).
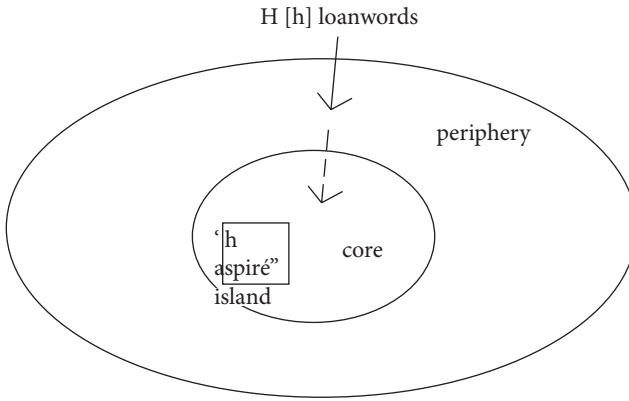

Figure 2. Friesner (2001) on the integration of 'h aspiré' loanwords in the French lexicon

Returning to the data, I find that consonant-final words are not a problem for stress assignment: they have final stress in Turkish as well as in Romanian, as shown in (16).

(16) Turk. susám 'sesame' > Rom. susán

Vowel-final words, on the other hand, must choose between maintenance of stress (17a,b) and expression of gender desinence (17c,d).

(17) a. Turk. pará 'money' > Rom. pará
      b. Turk. belá 'trouble' > Rom. beléą
      c. Turk. paʃá 'general' > Rom. páʃə
      d. Turk. kulé 'tower' > Rom. kúlə

After these considerations, only one example from the data remains unexplained, given in (18). Under the current assumptions, if the final vowel is adapted as a desinence vowel, the stress should move to the last syllable within the domain of stress. Instead, the outcome follows neither of the usual patterns observed of stress maintenance or stress shift to the nearest syllable.

(18) Turk. gøtyrý 'price/lump' > Rom. gjóturə 'a lot'

I offer only a speculative explanation for this result: it is possible that this outcome reflects secondary stress in the input from Turkish, which is maintained in Romanian as primary stress. I do not, however, purport to make any claims about the stress system of Turkish.

5.   Stress and desinence vowels in the French loanwords

One difference between French and Turkish that bears on gender assignment is that French nouns carry gender. French gender is usually maintained in the Romanian loan. In some cases, this gender assignment fits in nicely with the Romanian desinence classes. In other cases, more complications arise.

Easiest to adapt are French consonant-final nouns. In the feminine form, these nouns, which are often spelled with a final *e*, are sometimes pronounced with a final schwa in French. This equates nicely with the Romanian feminine desinence (19a,b). Consonant-final nouns also fit easily into the class of masculine nouns with underlying final -*u*, which does not surface (19c).

(19)   a.   Fren.   dentelle /dātél(ə)/ 'lace'   >   Rom.   dantélə
       b.   Fren.   étiquette /etikét(ə)/ 'label'   >   Rom.   etikétə
       c.   Fren.   boudoir /budwár/ 'boudoir'   >   Rom.   budwár

For vowel-final masculine nouns, the word's form is modified. This modification allows the maintenance of both stress position and gender desinence. Exceptionally, these nouns add a -*u* post-vocalically, which is realized as /w/. In the native vocabulary, the -*u* desinence generally follows a consonant. A couple of examples are given in (20).

(20)   a.   Fren.   pari /parí/ 'bet'   >   Rom.   paríw
       b.   Fren.   héros /eró/ 'hero'   >   Rom.   erów

The disparity here between French and Turkish cannot be explained by the time of borrowing, since earlier Turkish borrowings were sometimes not nativized. Instead, it seems likely that the difference is due both to a difference in the nature of contact and to linguistic differences between French and Turkish as compared to Romanian. French words already have a grammatical gender, which the initial borrowers may have been aware of and attempted to respect. The agents of borrowing from French were usually scholars, who would have learned French grammar formally and have a heightened awareness of words' gender because of having had to learn this specifically. There was also a need to have these words fit into a native pattern, so that they would seem more 'authentic', since French words were often borrowed out of a conscious effort to 're-Latinize' the language.

The few exceptions to the pattern in which French gender matches up with Romanian gender could reflect cases in which borrowers simply got the gender wrong. In these instances, such as those given in (21), the phonological form is likely to blame.

(21)   a.   Fren.   fantôme /fãtóm(ə)/ (*MASC.*) 'ghost'   >   Rom.   fantómə (*FEM.*)
       b.   Fren.   tournée /turné/ (*FEM.*) 'tour'   >   Rom.   turnéw (*MASC.*)

There are only a few other minor problems in the data set left to explain. First of all, there are a few cases in which final /u/ in French is maintained as a desinence within the domain of stress, while desinence vowels are generally excluded from this domain. Some examples are given in (22). This result seems to reflect a strong prohibition throughout the Romanian language against the sequence */uw/.

(22) a. Fren. acajou /akaʒú/ 'mahogany'   > Rom. akaʒú
      b. Fren. rendez-vous /rãdevú/ 'appointment' > Rom. randevú

There is one instance (23) in which final stressed /e/ is allowed, but this form varies with another, less exceptional form. In this case, according to Chitoran (p.c.), the form with final stressed /e/ is declined as if it contained final /ew/.

(23) a. Fren. purée /pyré/ 'purée' > Rom. pjuré/piréw

Finally, words that end in *-ie* (pronounced /i/) in French are problematic for Romanian. If only pronunciation is considered, these should be masculine in Romanian, but in French they are always feminine. A number of creative solutions have been devised in order to maintain stress and give some native-like gender desinence to such words. Examples are given in (24).

(24) a. Fren. jalousie /ʒaluzí/ 'Venetian blind' > Rom. ʒaluzęá
      b. Fren. galanterie /galãt(ə)rí/ 'gallantry' > Rom. galanteríe

In French, all words have final prominence, except that the optional final schwa is never stressed. In the Romanian loans, we observe little change in the position of stress in loanword adaptation. Unlike with the Turkish loanwords, maintenance of stress, in fact, takes precedence over maintenance of form, although in all cases a desinence vowel must still be present.

A possible formalism to account for the discrepancy between French and Turkish within Optimality Theory would be to rank DEP-IO below both REALIZE-MORPHEME(gender) *and* FAITH(stress) for this level of the vocabulary. This would capture the generalization that for French-origin loanwords, segment insertion is permissible in order to mark gender overtly. This also constitutes an allowable reranking of constraints for different strata of the lexicon under the assumptions of Davidson & Noyer (1997), since it requires only the reranking of a specific Faithfulness constraint.

Finally, examples such as (25) lend support for the stress pattern proposed here, in which desinence vowels fall outside the domain of stress. This example, the one preposition in the data set, exhibits final stress since, as a preposition, it does not require the addition of a desinence vowel.

(25) Fren. vis-à-vis /vizaví/ 'vis-à-vis' > Rom. vizaví

6. Conclusions

In this paper, I have demonstrated that many different factors affect the phonological shape of loanwords. Morphological factors constitute one important aspect that must be considered. Social factors are also relevant considerations.

While orthographic effects may be present in loanword adaptation, as has been demonstrated by Vendelin & Peperkamp (2006), they do not necessarily impede analysis. For example, final orthographic consonants that are not pronounced in French are occasionally realized in the Romanian loans, some examples of which are given in (26). However, this effect has no bearing on stress or gender desinence.

(26) a. Fren. blond /blɔ̃ / 'blond' > Rom. blónd
      b. Fren. boulevard /bul(ə)vár/ 'boulevard' > Rom. bulevárd

The findings presented here imply that only by contrasting borrowings from different languages into the same language can the relative effect of internal and external factors on the outcome of contact be determined. For example, French has a gender system and Turkish does not; this seems to call for an internal explanation for differences in treatment of gender. On the other hand, French and Turkish both have final stress; differences in stress assignment in the Romanian loanwords thus suggest an external explanation. Similarly, the fact that native Romanian words almost always carry a desinence vowel suggests an external explanation for the fact that Turkish loans do not always have a desinence vowel (and thus remain more foreign-sounding), while French loans almost always do (in order to appear more native-like).

Thus, in order to get a full picture, we must look for explanations within a number of areas of the language. These include different modules, such as the phonology and the morphology, as well as different levels, including linguistic differences and external explanations such as orthography and social factors.

References

Chitoran, Ioana. 2001. *The Phonology of Romanian: A Constraint-Based Approach*. Berlin: Mouton de Gruyter.

Close, Elizabeth. 1974. *The Development of Modern Rumanian*. London: Oxford University Press.

Davidson, Lisa & Rolf Noyer. 1997. "Loan Phonology in Huave: Nativization and the Ranking of Faithfulness Constraints." *Proceedings of WCCFL* 15. Palo Alto, Calif.: CSLI, 65–80.

Fisiak, Jacek. 1975. "Some Remarks Concerning the Noun Gender Assignment of Loanwords." *Bulletin de la Société polonaise de linguistique* 33.59–63.

Friesner, Michael L. 2001. *The Phonological Representation of French 'h aspiré' from the Perspective of Loanword Phonology*. B.A. Thesis, Dartmouth College, Hanover, N.H.

———— 2006. "Bracketing and Cyclicity in Romanian Stress." *University of Pennsylvania Working Papers in Linguistics* 12:1.95–108.

Harris, James W. 1991. "The Exponence of Gender in Spanish." *Linguistic Inquiry* 22.27–62.

Haugen, Einar. 1950. "The Analysis of Linguistic Borrowing." *Language* 26.210–231.

———— 1969. *The Norwegian Language in America.* Bloomington: University of Indiana Press.

Iscrulescu, Cristian. 2003. "Morphological Faithfulness and Phonological Markedness: The Case of Romanian Nominals." *USC Working Papers in Linguistics* 1.13–27.

Itô, Junko & R. Armin Mester. 1995a. "Japanese Phonology." *The Handbook of Phonological Theory*, ed. by John A. Goldsmith. Cambridge, Mass.: Blackwell, 817–838.

———— 1995b. "The Core-Periphery Structure of the Lexicon and Constraints on Reranking." *Papers in Optimality Theory. University of Massachusetts Occasional Papers* 18.181–210.

Kenstowicz, Michael. 2007. "Salience and Similarity in Loanword Adaptation: A Case Study from Fijian." *Language Sciences* 29.316–340.

———— & Hyang-Sook Sohn. 2001. "Accentual Adaptation in North Kyungsang Korean." *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz. Cambridge, Mass.: MIT Press, 239–270.

———— & Atiwong Suchato. 2006. "Issues in Loanword Adaptation: A Case Study from Thai." *Lingua* 116.921–949.

Maneca, C. 1966. "La struttura etimologica della lingua romena letteraria contemporanea dal punto di vista della frequenza dei vocaboli." *Revue roumaine de linguistique* 11:3.219–237.

Peperkamp, Sharon & Emmanuel Dupoux. 2002. "A Typological Study of Stress Deafness." *Laboratory Phonology VII*, ed. by Carlos Gussenhoven & Natasha Warner. Berlin: Mouton, 203–240.

Petrucci, Peter R. 1999. *Slavic Features in the History of Rumanian*. Munich: LINCOM Europa.

Poplack, Shana, Alicia Pousada & David Sankoff. 1982. "Competing Influences on Gender Assignment." *Lingua* 57.1–28.

Poplack, Shana, David Sankoff & Christopher Miller. 1988. "The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation." *Linguistics* 26.47–104.

Repetti, Lori. 2003. "Come i sostantivi inglesi diventano italiani: la morfologia e la fonologia dei prestiti." *Italiano e inglese a confronto*, ed. by Anna-Vera Sullam Calimani. Firenze: Franco Cesati, 31–42.

———— 2006. "The Emergence of Marked Structures in the Integration of Loans in Italian." *Historical Romance Linguistics: Retrospective and Perspectives*, ed. by Randall S. Gess & Deborah Arteaga. Amsterdam & Philadelphia: John Benjamins, 209–239.

Sala, Marius. 1976. *Contributions à la phonétique historique du roumain*. Paris: Klincksieck.

Suciu, Emil. 1992. "Quelques remarques sur l'adaptation phonétique des emprunts lexicaux roumains au turc-osmanli." *Revue des Études Sud-Est Européennes* 30.27–32.

Svensson, Ann-Marie. 2001. "Germanic vs. French Stress in Disyllabic Loan Words in English." *Studies in Honour of Göran Kjellmer. Gothenburg Studies in English* 81.123–129.

Thornton, Anna M. 2003. "L'assegnazione del genere ai prestiti inglesi in italiano." *Italiano e inglese a confronto*, ed. by Anna-Vera Sullam Calimani. Firenze: Franco Cesati, 57–86.

Underhill, Robert. 1976. *Turkish Grammar*. Cambridge, Mass.: MIT Press.

Vendelin, Inga & Sharon Peperkamp. 2006. "The Influence of Orthography on Loanword Adaptation." *Lingua* 116.996–1007.

Walker, Rachel. 2000. "Nasal Reduplication in Mbe Affixation." *Phonology* 17.65–115.

Weinreich, Uriel. 1968. *Languages in Contact*. The Hague: Mouton.

Mandarin adaptations of coda nasals in English loanwords*

Feng-fan Hsieh, Michael Kenstowicz & Xiaomin Mou
Massachusetts Institute of Technology

The paper documents and analyzes the ways in which English loanwords into Mandarin are adapted to conform to the Rhyme Harmony constraint that requires the front vs. back quality of a nonhigh vowel to agree with the coronal vs. dorsal character of a nasal coda. The principal finding is that the backness of the English vowel determines the outcome and can force a change in the place of articulation of the nasal coda. This is attributed to the phonetic salience of the vowel feature in comparison to the relative weakness of the nasal place feature. It is concluded that phonetic salience is a critical factor in loanword adaptation that can override a phonologically contrastive feature.

1.   Background and motivation

In the recent theoretical literature on loanword phonology two competing models have emerged. The first, championed by Paradis & LaCharité (1997, 2005) and others, holds that loanword adaptation is executed primarily by bilinguals who draw on their phonological competences in both the donor (L2) and recipient (L1) languages to discern segmental equivalences at an abstract, phonological (phonemic) level. When an exact phonemic match is not found then the closest available phoneme is chosen, with distance measured in terms of the distinctive features operative in the native, L1 grammar. An alternative view, typically couched within the OT model, sees loanword adaptation as based on the phonetic output of the donor language– either in the form of a raw acoustic signal (Silverman 1992) or more usually in a UG-based phonetic transcription of varying degrees of detail and abstraction.[1]

1.   Under standard conceptions, OT grammars lack an intermediate, phonemic (word-level) level of representation, making the kind of mapping envisioned by Paradis & LaCharité (1997, 2005) unavailable.

The adapter can take a variety of factors into account in order to make the loan sound like a word of the native language while still remaining as faithful as possible to the source of the loan. These include orthography as well as phonetic properties that are salient to an L1 speaker–regardless of their contrastive status in the L1 or L2 grammars. See Kenstowicz & Suchato (2006) and Yip (2006) as well as cited references for discussion of this alternative.

Mandarin Chinese presents us with the possibility of an interesting test of these two alternative models of loanword adaptation. According to most analyses (e.g. Duanmu 2000, 2007), Mandarin has five vowel phonemes: /i/, /y/, /u/, /ə/, and /a/. The high vowels contrast for [back] and [round] while the mid and low vowels do not. Stressed syllables are subject to a bimoraic constraint. There are no complex syllable margins. Codas are restricted to the nasals /n/ and /ŋ/ (modulo r-suffixation in the formation of the diminutive) and the glides/semivowels /j/, /w/. The canonical lexical item has the shape C(Gl)VX (X = V, Gl, N). The vowels take on a variety of allophonic guises depending on the surrounding consonants. In (1) we illustrate several generic CVV syllables. The first column is the Pinyin transliteration, the second is the underlying phonemicization, and the third is a broad phonetic transcription (Duanmu 2000).

(1) | Pinyin | UR | PR | |
|--------|----|----|-|
| tā | tʰā | tʰāa | 'she' |
| tí | tʰí | tʰʲíi | 'dam' |
| tǔ | tʰǔ | tʰʷǔu | 'mud' |
| tè | tʰə̀ | tʰ ɤ̀ɤ | 'special' |

In the context of nasal codas the low vowel takes a relatively front allophone before the dental nasal (typically transcribed as [an]) and a relatively back, unrounded allophone before the velar nasal (transcribed as [aŋ])–a distribution termed Rhyme Harmony in Duanmu (2000, 2007). By contrast, in English front and back low vowels freely combine with the dental and velar nasal phonemes to give four possible combinations.[2]

(2) | English | | Mandarin | | |
|---------|-|----------|-|-|
| [æn] | Dan | [an] | dan | 'egg' |
| [æŋ] | dang | | | |
| [an] | Don | [aŋ] | daŋ | 'swing' |
| [aŋ] | dong | | | |

If loanword adaptation abstracts away from the phonetic details in both L1 and L2 grammars, then we expect that in cases of conflict between faithfulness to the English vowel or to the nasal coda, the Mandarin adaptation should be determined

---

2. Before the velar nasal the vowel is rounded [ɒ] or [ɔ] for many English speakers.

by the nasal consonant. This is because the nasal coda is the only point of similarity at the phonological level, given that the vowel is unspecified or noncontrastive for [back] in Mandarin (indicated by the archiphoneme A; see Wang (1993) & Duanmu (2000, 2007) for details). This scenario is sketched in (3)

(3)   phonological mapping

English		Mandarin
æn		An
æŋ		
ɑn		Aŋ
ɑŋ		
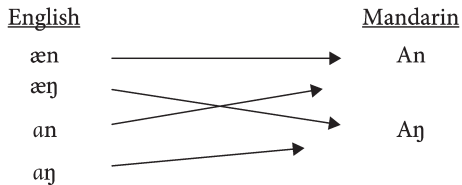
Alternatively, if the adapter is trying to achieve the best phonetic match then in cases of conflict (i.e. English [æŋ] and [ɑn]), additional considerations may come into play to decide the outcome. A priori we might expect variation across different lexical items depending on whether the vowel or the coda nasal is the determining factor. Alternatively, the adapter might call on other criteria to break the tie. For example, while the [±back] vowel difference is phonologically predictable, it is more salient phonetically and hence could provide a better overall match than the nasal coda consonant–a segment whose place features are relatively faint and highly susceptible to neutralization cross-linguistically. The latter scenario predicts the correspondences in (4).

(4)   phonetic (auditory) mapping

English		Mandarin
æn		an
æŋ		
ɑn		ɑŋ
ɑŋ		
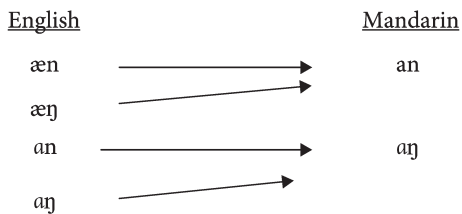
In the absence of a well-developed theory of loanword adaptation, it is unclear which of these two alternatives is more likely to be true. Hence the empirical study of such conflicting cases is an important step towards such a theory.

Whether (3) or (4) is the correct scenario turns out to be a question that is not so easily answered. It is well known that in comparison to Japanese and Korean, Mandarin Chinese is highly resistant to phonological loans, preferring loan translations or calques (Novotná 1967). Furthermore, it appears that many of the phonological loans that entered the language in the Early Modern period (c. 1900–1940) have been become obsolete or been replaced. Contemporary Mandarin vocabulary thus lacks a substantial body of loanwords that we can easily consult in order to

answer our question. We are thus forced to fall back on more meager resources. We are aware of two sources with relevant data. First, there is the *Dictionary of Loanwords and Hybrid Words in Chinese* (Liu et al. 1984).[3] We analyze material drawn from this source in Section 2. Second, there is the Website of the Chinese Ministry of Foreign Affairs, which has a listing of the preferred transcriptions and pronunciations for many foreign place names. We analyze data drawn from the latter in Section 3. Section 4 reviews the phonetic basis of the front-back vowel enhancement of the coda nasal contrast to provide independent support for our analysis. Section 5 is a brief summary and conclusion.

2. Analysis of loanwords from the dictionary

Our study's loanword corpus consists of c. 600 items drawn from Liu et al. (1984) that contain a VN sequence in the loan source (typically English). The discussion here focuses on items where the vowel of the source word is low or mid since this is where the vowel is phonologically unspecified or noncontrastive for front vs. back in Mandarin and the resolution of the conflict between faithfulness to the vowel vs. faithfulness to the coda nasal of the English word can be studied. We organize the data into several subcategories. The first consists of English VN rimes where V is nonhigh, N is a dental or velar nasal, and the syllable bears some degree of stress. Our main finding is that it is the front vs. back category of the vowel that determines the outcome. We then look at VN sequences drawn from final unstressed syllables in English. Here we find competition between strategies based on approximation to the English reduced vowel vs. those based on the orthographic representation. The next category consists of loans in which a nasal has been inserted into the coda to achieve a bimoraic syllable. Our data indicate that the front vs. back quality of the vowel in English determines the substitution as [n] or [ŋ], respectively. In the last group, the coda nasal of the English loanword is [m], which must be repaired–typically by changing the [m] to [n] or [ŋ]. Once again, we find that the vowel of the source word decides the outcome.

2.1 VN

When the vowel is low there is a partial correlation between its front vs. back status in the English source word and its orthographic form. Front vowel [æ] (RP [a]) is

---

3. The authors state that the dictionary was constructed in the period 1960–64 from material in dictionaries, monographs, Chinese translations of foreign classics, academic journals, newspapers, magazines, as well as other sources such as import-export catalogs, customs declaration forms, etc. The dictionary contains c. 10,000 items.

represented with the letter *a* (e.g. *hat*) while the back, unrounded [ɑ] (or rounded [ɒ] in RP) is typically represented with *o* (e.g. *hot*). However, sometimes [ɑ] is also spelled with the letter *a* (e.g. *class*). Since we cannot always rely on English spelling, we have checked all examples with the OED.[4]

When the English source consists of a front vowel combined with a dental nasal ([æn]) or a back vowel combined with a velar nasal ([ɑŋ]), we expect the Mandarin adaptation to contain a matching rime–i.e. [an] or [aŋ], respectively. For English [æn] there are 31 loans in our corpus; all but five support this hypothesis.

(5)   <u>English</u>   <u>Mandarin</u>

a.   *an*chovy   [æn]   *an*.chou [an]
     *an*gel   *an*.qi.er
     *an*tecedent   *an*.ti.xi.deng
     fla*n*ge   fa.*lan*.(-pan)
     V*an*dyke   *fan*.tai.ke
     v*an* de graaf   *fan*.de.ge.la.fu
     fur*an*   fu.r*an*
     c*an*delilla   k*an*.te.li.la
     cl*an*   ke.*lan*
     cot*an*gent   kou.t*an*.jin
     l*an*cers   l*an*.sa.si
     r*an*d   l*an*.te
     l*an*thanium   l*an*
     Lev*an*t   li.*fan*.de
     rom*an*tic   luo.m*an*.di.ke
     rom*an*ce   luo.m*an*.si
     m*an*tle   m*an*.tuo
     p*an*dora   p*an*.duo.la
     sarab*an*d   sa.la.b*an*.de
     s*an*dal   sh*an*.da.li
     Sud*an*   su.d*an*
     t*an*gent   t*an*.jin

b.   b*an*dage   [æn]   b*eng*.dai   [ɤŋ]
     phal*an*stery   fa.*lang*.ji.si.te   [aŋ]
     sc*an*dium   *kang*
     l*an*tum   *lang*.tang
     v*an*dal   *wang*.da.er

For English [ɑŋ] rimes there are seven examples in the corpus; five are adapted as expected with a velar nasal and back vowel allophone (6a).

---

(6)   English   Mandarin

a.   *Con*go   [aŋ]   *gang*.guo   [aŋ]
      fra*nc*   fa.*lang*
      fur*long*   *lang*
      ping*pong*   ping.*pang*
      *mon*goose   *meng*.ge   [ɤŋ]

b.   *en*core   [aŋ]   *an*.ge   [an]
      gong   *gun*.ge   [un]

The matrix in (7) summarizes the adaptation of the harmonic rhymes. Mandarin preserves the front vs. back quality of the rhyme to a significant degree.

(7)   English

	an	aŋ
æn	26	4
aŋ	2	5

Mandarin $p < 0.008$ (two-tailed Fisher's exact test)

In loans where the English vowel and coda nasal do not agree as front vs. back, there are two ways in which the adaptation can be brought into alignment with the Mandarin [an] and [aŋ] codas required by Rhyme Harmony. Either the front vs. back character of the vowel can be preserved and the nasal changed; or alternatively the nasal coda can be held constant and the vowel adjusted. The data overwhelmingly evidence the first strategy. The corpus contains 24 loans where the English source consists of a low, back vowel and a dental nasal. In all of the corresponding Mandarin loans, it is the nasal consonant that is changed, giving an [aŋ] rhyme in the majority of cases (8a). In a few (8b), the vowel is mid [oŋ] or [ɤŋ].

(8)   English   Mandarin

a.   an*on*(ym) [an]   a.*nang*   [aŋ]
      *oun*ce   *ang*.si
      Br*own*ing   bai.*lang*.ning
      p*oun*d   *bang*
      bez*ant*   bie.*sang*
      rad*on*   d*ong*
      Oreg*on*   e.le.g*ang*
      erg*on*   er.g*ang*
      c*on*cept   g*ong*.si.bu.tuo
      fa*n*dango (Sp.)   *fang*.dange

	ge*on*	ji.*ang*.ding.sheng	
	co*n*denser	k*ang*	
	ca*n*to (It.)	k*ang*.tuo	
	cr*ow*n	ke.l*ang*	
	marath*on*	ma.la.s*ong*	
	m*on*soon	m*ang*.xun	
	p*on*tificate	p*ang*.ti.fei.jia.te	
	p*on*toon	p*ang*.tong	
	pla*n*te (Fr.)	pu.l*ang*.te	
	sa*m*ba	s*ang*.ba	
	so*n*net	sh*ang*.lai-ti	

    b.   go*n*dola   [an]   g*ong*.duo.la   [oŋ]

	ne*on*	ni.h*ong*	
	cell*on*	se.l*ong*	
	m*on*t	m*eng*	[ɤŋ]

The number of loanwords with a velar nasal coda and front vowel nucleus is again smaller (13).[5] Only four remain faithful to the nasal (9b). The rest (9a) change the [ŋ] to [n] so as to remain faithful to the English vowel.

(9)   <u>English</u>   <u>Mandarin</u>

    a.   ba*nk*   [æŋ]   ba*n*.ke   [an]

	*An*gora	*an*.ge.la	
	Fra*n*klin	fu.l*an*	
	Gra*n*ge	ge.l*an*.qi	
	La*n*cashire	l*an*.kai.xia	
	La*n*gley	l*an*.le	
	ta*n*go	t*an*.ge	
	ta*n*k	t*an*.ke	
	tria*n*gle	te.li.*an*.ge.er	

    b.   ga*ng*sa   [æŋ]   g*ang*.sha [aŋ]

	sara*ng*i	sa.l*ang*.ji	
	wa*n*kel	w*ang*.ke.er	
	Ya*n*kee	y*ang*.ji	

The matrix in (10) summarizes the resolution of the conflicting English rimes.

---

5.   The word *mu.si.deng* [dɤŋ] < *mustang* [æŋ] has an unexpected mid rather than low vowel. We are not able to explain this change in height. Given that it is treated as mid, the expected [den] syllable is rare in Modern Mandarin and is avoided in loans. See discussion of mid vowels below. The *Yankee* > *yang.ji* loan might arise from semantic contamination since the character used to represent it means 'western'.

(10)   English

	an	aŋ
æŋ	9	4
an	0	24

Mandarin   p < 0.000001 (two-tailed)

In sum, our hypothesis is supported–the more salient vowel normally determines the adaptation even though the nasal coda is the site of the phonological contrast.

Adopting the approach to loanword phonology taken in Kenstowicz (2005) and Yip (2006) where faithfulness to the loanword source is expressed as an OT Output-Output faithfulness constraint that may be ranked differently from the corresponding Input-Output constraint of native grammar, we can express the adaptation of the low vowel+nasal coda words into Mandarin as follows. First, we assume an undominated markedness constraint of RHYME HARMONY (Duanmu 2000, 2007) that requires a front vs. back low vowel to co-occur with a dental vs. velar nasal coda, respectively (see Flemming 2003 for discussion of the phonetic basis for such a constraint). Second, we assume that the nasal codas are the site of the lexical contrast in Mandarin (F » M) while the [±back] low vowel allophones [a] and [a] are distributed by Rhyme Harmony (M » F). Given the OT premise of Richness of the Base, native grammar inputs in which the nucleus and coda violate Rhyme Harmony are repaired by faithfulness to the coda, as in (11) below.

(11)

/an/	Rhyme Harmony	Id-CPl-Coda	Id-[back]
[an]	*!		
>[an]			*
[aŋ]		*!	
/aŋ/			
[aŋ]	*!		
>[aŋ]			*
[an]		*!	

But in the loanword phonology, in order to be faithful to the vowel of the source language, the adapter calls on the otherwise submerged Id-[back] constraint which is "cloned" as an Output-Output constraint between English and Mandarin and ranked above Faithfulness to CPl-Coda.

(12)   Id-[back]$_{E\text{-}M}$ »   Id-CPl-Coda   »   Id-[back]

Given this ranking, the input-output adaptation mapping is diverted towards faithfulness to the otherwise redundant vowel, as shown below.

(13)

/ɑn/	Rhyme Harmony	Id-[back]$_{E\text{-}M}$	Id-CPl-Coda
[ɑn]	*!		
[an]		*!	
>[aŋ]			*
/æŋ/			
[ɑŋ]	*!		
[aŋ]		*!	
>[an]			*

Let us now consider examples where the English source word consists of a mid vowel followed by a nasal coda. In Mandarin there are four surface mid vowels whose distribution is determined by the surrounding onset and coda consonants (Duanmu 2000, 2007). The basic allophone, found in open syllables, is back unround [ɤ]. As with the low vowel, a dental nasal requires a more fronted vowel nucleus [ən] (e.g. *sen* [sən] 'forest') while a velar nasal requires a back vowel nucleus [ɤ] or [o]. In some varieties of Mandarin the latter derives from earlier [uŋ] by lowering (e.g. [tóŋ] 'same'). Dialects also differ in whether or not [oŋ] is retained after a labial onset: cf. Taiwanese Mandarin [mǒŋ] 'fierce' vs. Beijing [mɤ̌ŋ]. Finally, there is a more close front vowel allophone after a palatal glide onset [je].[6]

---

6.   In order to have some sense of the location of these allophones in phonetic space, we recorded five tokens for each from a male Taiwanese Mandarin speaker (the first author). The results showing the average first and second formant measures and standard deviations from the mid point of the vowel are shown below. We see that the [ə] is a relatively central vowel

Turning to the loanwords, there are three cases to consider depending on whether the English vowel is [ɛ], [o], or [ʌ]. We examine each of these in turn. First, when the mid vowel is [ɛ] and the coda is [n] in the English source, Mandarin offers the choice between [je] and [ən]. Neither option is particularly close. It is therefore of some interest that the former is systematically rejected in favor of the latter (14).

(14)   English   Mandarin

     amen   [ɛn]   a.men   [ən]
     pentyl   ben.ti.er
     benzene   ben
     benzocaine   ben.zuo.ka.yin
     Enfield   en.fei.er
     phen(ol)   fen
     phosgene   fu.su.zhen
     convention   kang.wen.xin
     pimento   pi.men.ta
     cement   shui.men.ting

The choice of central [ə] over diphthongal [je] for English [ɛ] indicates that Dep-Glide dominates faithfulness for [back].

(15)

/bɛn/	Dep-Gl	Id-[back]$_{E-M}$
[bən]		*
[bjen]	*!	

The few exceptions to this correspondence occur when the C[ən] syllable is either not attested in the existing inventory of Mandarin syllable types or is rare (16). In this case an adjustment must be made–changing the vowel or the nasal.[7]

---

falling roughly midway between the front [je] and the back rounded [o] in F2. The nucleus of the [je] is more close, showing the influence of the onglide.

		F1	F2	
bēn	[bān]	472/9	1476/32	'run'
bēng	[bɤ̃ŋ]	471/12	1100/68	'collapse'
sōng	[sōŋ]	465/20	857/47	'loose'
bjēn	[bjēn]	366/15	2352/31	'edge'

7.   Since the data in the dictionary are all transcribed with Chinese characters, a syllable containing a novel combination of CVC cannot be easily represented. It is not clear to us to what extent this fact about orthography inhibits the creation of novel combinations of onset, rhyme, and tone. See Bauer (1985) for novel syllables in loans to Cantonese.

(16) Ländler (Germ) li*an*.de.la *[lən]
      l*en*tor   l*un*.tuo   *[lən]
      *en*gine   y*in*.qing   *[ən]
      t*en*dency   t*ing*.deng.se *[tən], *[tin]

Curiously, it is the vowel height that is changed while the front vs. back property of the vowel in the source word is largely maintained. Lin (2008) reports a similar finding. This suggests that faithfulness for the vowel is broken down into faithfulness for [±back] (F2) vs. faithfulness for height [±high], [±low] (F1), as indicated in the tableau in (17). (We assume the undominated constraint USE-LISTED-SYLLABLE: A syllable in the adapted loanword must have a precedent in the native inventory).

(17)

/ten/	Use-Ld-Syll	Id-[back]$_{E-M}$	Id-[high]$_{E-M}$	Id-CPl-Coda
ten	*!			
tin	*!			
>tiŋ			*	*
toŋ		*!		*

Our corpus contains 12 examples of conflicting English rimes in [on]. In the corresponding Mandarin adaptations, they change the [n] to velar [ŋ] in order to remain faithful to the vowel (18a). The lone exception is shown in (18b). When the onset is a labial consonant then the vowel [o] is blocked by the labial disharmony constraint that bans the combination of a labial onset and rhyme in the Beijing dialect. The back unrounded vowel [ɤ] (Pinyin *eng*) or the low [a] is substituted in this case.

(18)   <u>English</u>   <u>Mandarin</u>
     a. amid*one*   [on]   a.mi.t*ong*   [oŋ]
         barbit*one*   ba.bi.t*ong*
         chal*one*   ka.er.l*ong*
         cl*one*   ke.l*ong*
         Col*ogne*   ke.l*ong*
         horm*one*   he.er.m*eng* [ɤŋ]
         teleph*one*   de.lu.f*eng*
         microph*one*   mai.ke.f*eng*
         sousaph*one*   su.sha.f*eng*
         sarrusoph*one*   sa.luo.suo.f*eng*
         le*one*   li.*ang*   [aŋ]
     b. sc*one*   [on]   shi.g*an*   [an]

The contingency table in (19) summarizes the outcome of the competing changes for the [±back] feature of the vowel and the corresponding coronal vs. dorsal place feature of the nasal coda.

(19)   English

	Vn	Vŋ
en	14	2
on	1	11

Mandarin   p = 0.000053 (two-tailed)

As the tableau in (20) shows, in the case of the conflicting back vowel + coronal coda the correct adaptation is made by the Id-[back]$_{E-M}$ » Id-CPl-Coda ranking established for the low vowels in (12).

(20)

/on/	Rhyme-Harmony	Id-[back]$_{E-M}$	Id-CPl-Coda
[on]	*!		
> [oŋ]			*
[ən]		*!	

The behavior of the English rimes composed of the centralized, wedge vowel [ʌ] suggests that it is not salient enough on the crucial [±back] F2 dimension to force a change in the nasal coda. Faithfulness to the coda obtains in all but one case (21). The vowel receives a range of adaptations as high, mid, or low.

(21)   English   Mandarin
     a.   *un*cial   [ʌn]   *an*.se.er   [an]
          p*un*ch   p*an*.qu
          h*un*dredweight   h*an*.jue.huai.tuo
          carbor*un*dum   ka.bo.*lan*.deng
          Br*un*swick   bu.*lun*.si.wei.ke   [un]
          s*un*dae   sh*eng*.dai   [ɤŋ]
     b.   Y*oung*   [ʌŋ]   *yang*   [aŋ]

To summarize the adaptation of English VN rimes with a mid vowel, we find [ɛn] and [on] primarily rendered as faithful to vowel quality at the expense of change in the nasal consonant. The adaptation of [ʌn] and [ʌŋ] is determined by the nasal

coda, indicating that the wedge vowel is not decisive and reflecting its intermediate position on the [±back] F2 dimension.

Before turning to reduced vowels, we note a minor pattern. Seven items in our corpus terminate in the graph -*oon*. Since [dun], [tun], [sun] and [lun] are valid Mandarin syllables, it is puzzling why these adaptations are rejected in five of the seven items–primarily in favor of C[ong]. Possibly in these cases the adapter was following a graphic rule that interprets the -*oon* as if it terminated in a tense mid vowel, (i.e. [on]) rather than the phonetic [un].

(22) English   Mandarin

      card*oon*   [un]   ha.d*un*   [un]
      mons*oon*   mang.x*un*   [yn]
      cart*oon*   ka.t*ong*   [oŋ]
      pantal*oon*   pa.ta.l*ong*
      pont*oon*   pang.t*ong*
      shall*oon*   xia.l*ong*
      sim*oon*   xi.m*eng*   [ɤŋ]

English loans with final syllables containing unstressed, reduced vowels exhibit two competing adaptation strategies. The primary one substitutes *en* [ən]–arguably Mandarin's best phonetic match to the schwa-like, reduced vowel of English (23a). This practice is followed unless an illegal or rare syllable such as *len* [lən] or *den* [tən] results, in which case a high vowel is typically substituted (23b) instead. In a few cases (23c), the adapter has based the choice on the spelling.

(23)   English   Mandarin

     a. Addision   [ən]   a.di.s*en*   [ən]
         eikonog*en*   ai.ke.nu.zh*en*
         Brem*en*   bu.le.m*en*
         predicam*ent*   bu.li.di.jia.m*en*
         cushi*on*   gu.ch*en*
         clais*en*   ke.lai.s*en*
         co.se.c*ant*   kou.xi.g*en*
         li.n*en*   lian.r*en*
         mamm*on*   ma.m*en*
         Morm*on*   mo.m*en*
         bac*on*   pei.g*en*
         pullm*an*   pu.er.m*en*
         salm*on*   sa.m*en*
         cinchoph*en*   xin.ke.f*en*
         uni*on*   yu.r*en*

     b. Applet*on* [ən]   a.pu.d*un*   [un]   [dən] (rare)
         dal.t*on*   dao.er.d*un*
         west*on*   wei.si.d*un*

	bar*on*	ba.*lun*	*lən
	per.l*on*	bei.*lun*	
	gall*on*	jia.*lun*	
	carr*on*	ka.*lun*	
	Corbill*on*	kao.bi.*lun*	
	chaldr*on*	qiao.te.*lun*	
c.	sat*an* [ən]	sa.d*an* [an]	
	ti.t*an*	tai.t*an*	
	Zi*on*	xi.*an*	
	harmatt*an*	ha.ma.d*an*	

These data can be analyzed by assuming that the schwa of the source word is not salient enough to determine the outcome and the decision is passed on to the coda nasal. Since all the examples have coda [n], no change is required. Our corpus contains no unstressed syllables with a coda dorsal nasal, which in any case are rare with nonhigh vowels in English. The tableau below illustrates the adaptation *baron* > *ba.lun*. The adaptations with the high vowel in (23b) indicate that [+high] is preferred to [+low] as a match for the unstressed schwa of English, probably because high vowels are phonetically shorter than low vowels cross-linguistically.

(24)

/bærən/	Use-Ld-Syll	Id-[back]	Id-CPl-Coda	Id-duration
lən	*!			
loŋ		*	*!	
lan		*		*!
> lun		*		

In sum, the adaptations analyzed in this section indicate that when the vowel of the English source word is front or back then it determines the way in which the loan accommodates the Rhyme Harmony constraint. Nonsalient schwa or wedge seem to pass the decision on to the nasal coda.

In the next two sections we review a couple of other places in the loanword grammar where the place feature of a nasal coda is determined by the vowel of the source word.

2.2   V.NV → VN.NV

In (25) we list examples in which a nasal consonant is added to the coda before a following nasal onset in order to satisfy the bimoraic requirement on stressed

syllables. Interestingly, the choice between [n] and [ŋ] is determined, not by gemi-
nating the nasal of the source word, but rather by the vowel of the augmented syl-
lable (25a). For example, in the adaptation of *economy* the English stressed syllable
is augmented in the Mandarin loan by insertion of a velar rather than a dental
nasal: *ai.kang.nuo.mi.* The handful of exceptions to this generalization is shown
in (25b).[8]

(25)   English   Mandarin

   a.   am*o*n*al   [on]   a.m*ang*.na   [aŋ]
        ec*o*n*omy   [an]   ai.*kang*.nuo.mi [aŋ]
        *an*archy   [æn]   *an*.na.qi   [an]
        be*n*adryl   [en]   b*en*.na.jun   [en]
        fel*o*ny   [ən]   fei.*lun*.nu   [un]
        laud*a*nidine   [æn]   lan.d*an*.ni   [an]
        m*a*n*a   [æn]   m*an*.na   [an]
        m*o*narchy   [an]   m*eng*.ne.a.ji   [ɤŋ]
        perph*e*nazine   [en]   pia.f*en*.na.xin [en]
        pe*n*icillin   [en]   p*an*.ni.xi.lin [an]
        thir*am*   [æm]   qiu.*lan*.mu   [an]
        se*n*eca   [en]   s*en*.ni.jia   [en]
        pe*n*ny   [en]   p*en*.ni   [en]
        arsph*e*namine   [en]   shen.f*an*.na.ming   [an]
        sp*a*nner   [æn]   shi.b*an*.na   [an]
        sc*am*mony   [æn]   si.*kan*.mo.ni   [an]
        T*am*many   [æm]   t*an*.mu.ni   [an]
        gu*n*ny   [ʌn]   g*ong*.ni   [oŋ]
        T*o*ny   [on]   t*ang*.ni   [aŋ]
        Down*i*ng   [aun]   t*ang*.ning   [aŋ]

   b.   afgh*a*ni   [an]   a.fu.h*an*.ni   [an]
        m*e*mory   [em]   m*eng*.mo.li   [ɤŋ]
        m*am*moth   [æm]   m*eng*.ma   [ɤŋ]

The data in (25) show that faithfulness to the backness of the vowel–redundant
in Mandarin but contrastive in English–is an active constraint of the loanword
grammar that overrides homorganicity for the NC cluster that might otherwise be
expected since it does not require the insertion of a place feature in the coda but

---

8.   The OED indicates a back vowel for the medial syllable of *Afghani* in (25b). For this reason
we classify it as an exception. Also the loan *mammoth > meng.ma* is represented with the char-
acter for *meng* 'fierce', perhaps for semantic reasons.

merely anticipates the place feature of the following onset. We illustrate this aspect of the Mandarin loanword grammar in (26).

(26)

e/kɑn/omy	Id-V[back]$_{E-M}$	Dep-Place-Coda
ai.kan.nuo.mi	*!	
>ai.kɑŋ.nuo.mi		*

2.3　Vm → Vn, Vŋ

Finally, in (27) we list loans where the English source word contains a labial nasal in the coda. Since Mandarin bars [m] from the coda, the nasal coda must alter its place of articulation. The data indicate that the choice between [n] and [ŋ] is determined primarily by the front vs. back nature of the preceding vowel in the English source word (27a). The more centralized wedge vowel is once again less decisive, occurring with both dorsal and well as coronal nasal codas. Exceptions are shown in (27b).[9] Here as well we find that the adaptation has recourse to the more salient vowel rather than substituting a default consonant such as [n] that might otherwise be expected under a *Dorsal » *Labial » *Coronal ranking for consonantal place (de Lacy 2006).

(27)　　　　English　　　　　　　Mandarin

　　　a.　　*am*bersite　　[æm]　　*an*.bu.rui.te　　[an]
　　　　　　*am*broise　　[æm]　　*an*.bu.luo.si
　　　　　　*am*pul　　[æm]　　*an*.bu
　　　　　　s*am*sonite　　[æm]　　s*an*.suo.na.te
　　　　　　Gr*am*　　[æm]　　ge.l*an*
　　　　　　j*am*　　[æm]　　zh*an*
　　　　　　c*om*post　　[am]　　k*ang*.po.si.te　　[aŋ]
　　　　　　c*om*bination　　[am]　　k*ang*.bai.na.xiong
　　　　　　c*om*post　　[am]　　k*ang*.po.si.te
　　　　　　c*om*pote　　[am]　　k*ang*.bo.te
　　　　　　c*om*munism　　[am]　　k*ang*.men.ni.si.mu
　　　　　　c*om*mons　　[am]　　k*ang*.men.si
　　　　　　c*om*biner　　[am]　　k*ang*.ping.na
　　　　　　c*om*mission　　[am]　　k*ang*.mi.xiong

9.　The *shampoo* > *xiang.bo* loan is represented with the character for 'fragrance' and so may be a case of semantic contamination.

	Tom	[am]	tang.mu	
	Thompson	[am]	tang.mu.sheng	
	samba	[am]	sang.ba	
	quinoform	[om]	kui.nuo.fang	
	embelin	[em]	en.bei.lin	[ən]
	sumbul	[ʌm]	sang.bo	[aŋ]
	gumbo	[ʌm]	gong.bo	[oŋ]
	rum	[ʌm]	lan.mu	[an]
	calumba	[ʌm]	ka.lun.ba	[un]
	yumpies	[ʌm]	yong.pi.si	[oŋ]
	carborundum	[ʌm]	ka.bo.lan.deng	[ɤŋ]
	trumpet	[ʌm]	qu.lang.pai.ti	[aŋ]
	atom	[əm]	a.tun	[un]
	Edam	[əm]	yi.dun	[un]
b.	shampoo	[æm]	xiang.bo	[aŋ]
	mambo	[am]	man.bo	[an]
	adam	[əm]	ya.dang	[aŋ]
	empire	[em]	ying.bai.er	[iŋ]
	emperor	[em]	ying.bai.li.re.er	[iŋ]

The tableau in (28) shows the effect of Ident-V[back]$_{E-M}$ in the adaptation of *compost*.

(28)

/kɑmpost/	Id-V[back]$_{E-M}$	*dorsal, labial	*coronal
kan.po.si.te	*!		
> kaŋ.po.si.te		*	

The adaptations of the unstressed syllables of *atom* and *Edam* ([atun] and [yidun]) with a coronal support the idea that [n] is the default nasal. If the [±back] quality of the schwa vowel of English is indeterminant (as seems plausible–cf. Flemming & Johnson 2007) then the choice between the coronal and dorsal coda is resolved by the markedness hierarchy that substitutes coronal as the default oral place.

(29)

/ætəm/	Id-V[back]$_{E-M}$	*dorsal	*coronal
>atun			*
atuŋ		*!	

Before concluding this section we briefly address the possible role of orthography in the coda nasal adaptation process. The vast majority of loans spelled with "on" are adapted as [aŋ] and those spelled with "an" as [an]. Could orthography be the basis of the adaptation pattern rather than reference to the salience of the vowel on the F2 dimension? We think not. First, the few words in our corpus with an [an] sequence from Romance languages such as *franc* and *canto* are adapted with a back rhyme in accord with the back vowel in the source. Second, "on" and "an" words where the corresponding syllable in English is unstressed such as *cushion* are by and large adapted with [ən] and not [aŋ] or [oŋ]. Since stress is not orthographically recorded, the adaptation must be based on the spoken form of the word to explain this distinction. In our view the adapters use their knowledge of the spelling regularities of the source languages to guide them in the correct pronunciation of the source word vowel, which in turn determines the adaptation. This is evident in occasional mistaken interpretations such as *satan* > [sa.dan] where the final syllable is treated as stressed. Finally, even if we were to grant that orthography is the basis of the adaptation, it does not help to explain why in the orthographic equivalence of "on" = [aŋ], it is the vowel symbol that is the determining factor rather than the consonantal one. Salience in the phonetics provides a more plausible basis for understanding the adaptation, especially when it is combined with the observation that the less salient central vowels wedge and schwa do not determine the outcome. Here it is the nasal consonant that appears to do so. If the Mandarin adaptation of nasal codas is based on spelling, then if "n" determines the outcome for syllables with wedge and schwa, why not for "an" and "on" as well?

3.   Another corpus

The list of place names on the Chinese Ministry of Foreign Affairs website provides another opportunity to study the adaptation of nasals into Mandarin.[10] These data largely corroborate the generalizations found in the data from the dictionary discussed in Section 2. First, source words with a front [æ] followed by a nasal ([m] or [n]) uniformly have the [an] correspondence in Mandarin.

(30)   <u>English</u>   <u>Mandarin</u>
       Alex*an*der   [æn]   ya.li.sh*an*.da   [an]
       Amsterd*am*   a.mu.si.te.d*an*
       *An*atolia   *an*.na.tuo.li.ya
       Atl*an*ta   ya.te.l*an*.da

---

10.   We thank Ross Foo for providing us with a transcribed list of such words.

Birmingh*am*	bo.ming.h*an*
C*an*berra	k*an*.pei.la
C*an*terbury	k*an*.te.bo.lei
C*an*ton	k*an*.dun
F*an*ning	f*an*.ning
Fl*an*ders	fo.l*an*.de
Gr*am*pian	ge.l*an*.ping
H*am*pshire	h*an*.pu
Indi*an*a	yin.di.*an*.na
K*an*sas	kan.sai.shi
M*an*chester	m*an*.che.si.te
M*an*hattan	m*an*.ha.dun
M*an*sfield	m*an*.si.fe.ier.de
N*an*cy	n*an*.xi
N*an*tucket	n*an*.ta.ji.te
Nottingh*am*	nuo.ding.h*an*
St*am*ford	si.t*an*.fu

In cases where the loan source contains a conflicting combination of vowel nucleus and nasal coda, the vowel is the determining factor in the adaptation in the vast majority of cases (31). For [an] *Ontario* and *Tucson* (31b) are the exceptions where we find an unexpected front vowel. Perhaps the latter is based on a false parsing *Tuc+son* (cf. *Addison > adisen*).For [æn] the only exception is *Doncaster* (31d), for which the OED provides a [dæn] transcription despite the spelling.

(31)

	English		Mandarin	
a.	Adir*on*dacks	[an]	a.di.*lang*.da.ke	[aŋ]
	Br*on*x		bu.*lang*.ke.si	
	C*on*nacht		k*ang*.nuo.te	
	C*orn*wall		k*ang*.wo.er	
	Kl*on*dike		ke.*lang*.dai.ke	
	Oreg*on*		e.le.g*ang*	
	Wisc*on*sin		wei.si.*kang*.xing	
	L*ong*ford		*lang*.fu.de	
	T*aun*ton		t*ang*.dun	
	Tyr*one*		di.l*ong*	[oŋ]
	Yuk*on*		yu.k*ong*	
	M*on*tana		m*eng*.da.na	[ɤŋ]
	M*on*te Carlo		m*eng*.te.ka.luo	
	M*on*tpelier		m*eng*.bi.liai	
	Verm*ont*		fo.m*eng*.te	
b.	*On*tario	[an]	*an*.da.lue	[an]
	Tucs*on*		tu.s*en*	[ən]
	P*on*dicherryu		b*en*.di.zhi.li	

c.	*An*chorage	[æŋ]	*an*.ke.lei.qi	[an]
	*An*guilla		*an*.gui.la	
	*An*gus		*an*.ge.si	
	Fra*nk*furt		fa.l*an*.ke.fu	
	Fra*nk*lin		fu.l*an*.ke.lin	
	La*nc*ashire		l*an*.kai.xia	
d.	D*on*caster	[æŋ]	t*ang*.ke.si.te	[aŋ]

We have seven examples for [ʌŋ]. Five are faithful to the nasal, recapitulating the behavior seen earlier in (21). The syllable gaps *len and *den motivate the changes in vowel height.

(32)

English		Mandarin		
F*u*ndy	[ʌn]	f*en*.di	[en]	
Br*u*nswick		bu.l*un*.rui.ke	[un]	*len
L*o*ndon		l*un*.d*un*		*len, *den
D*u*nkirk		d*un*.ke.er.ke		*den
Fr*o*nt		fu.l*an*.te	[an]	*len
S*u*nderland		s*ang*.de.lan	[aŋ]	
D*u*ndee		d*eng*.di	[ɤŋ]	*den, *din

Finally, when the final syllable in the English loan is unstressed, the expected [Cən] is found in many cases (33a). (33b) and (33c) reflect two alternative responses to the Mandarin syllable gaps against the otherwise expected [ən] rime. In the former the vowel is adapted as high (Id-CPl-Coda » Id-V[high]) while in the latter the height change is blocked, compelling a change of the nasal coda (Id-V[high] » Id-CPl-Coda). The adaptations in (27d) appear to be based on the orthographic representations in which the vowels are treated as full rather than reduced.

(33)

	English		Mandarin		
a.	Cardig*an*(shire)	[ən]	ka.di.g*en*	[ən]	
	Dev*on*		de.w*en*		
	Linc*oln*(shire)		lin.k*en*		
	Log*an*		luo.g*en*		
	New Hav*en*		niu.hei.w*en*		
	Sax*ony*		sa.kr.s*en*		
	Solom*on*		suo.luo.m*en*		
b.	Bost*on*	[ən]	bo.shi.d*un*	[un]	*den
	Et*on*		yi.d*un*		*den
	Lachl*an*		la.ke.l*un*		*len
	Lawr*en*ce		lao.l*un*.si		*len

c.   Croyd*en*   ke.luo.yi.d*eng*   [ɤŋ] *den
     Wimbled*on*   wen.bu.er.d*eng*   *den

d.   Akr*on*   [ən]   a.ke.l*ong*   [oŋ]
     Cro-Magn*on*   ke.lu.mai.n*ong*
     Birmingh*am*   bo.ming.h*an*   [an]
     Ev*an*s   ai.*fan*
     Michig*an*   mi.zhi.*an*
     New Orle*an*s   niu.ao.er.li*ang* [aŋ]

In sum, the adaptations in the place names largely conform to the generalizations uncovered in the analysis of the dictionary loans in Section 2.[11] The adaptation of the coda nasal is determined by the position of the vowel in the source word on the front-back, F2 dimension. When the vowel occupies an intermediate position on this dimension, as in the case of wedge [ʌ], or is indeterminant, as in the case of schwa, the nasal place of the coda is largely preserved.

4.   Phonetic basis

The surface phonetic contrast in vowels before the alveolar vs. dorsal codas has been studied in a number of phonetic investigations of Mandarin. For example, Chen (2000) reports F2 differences of c. 500 Hz. in [an] vs. [aŋ] rhymes when they appear before a stop such as dān.dǎ 'single hit' (tennis) vs. fàn.dà 'magnify'. They are located in the interior of the vowel and are not just a coarticulatory effect at the VC transition. Crucially, she also finds that these differences persist–at a lower (c. 250 Hz) but still significant (P < 0.001) magnitude–in the wake of the deletion of an intervocalic nasal in casual speech *shā(n).ào* 'cove'. The magnitude of the F2 differences in Mandarin [an] vs. [aŋ] rhymes was further documented by Mou (2006), who found a c. 400 Hz difference at the mid point of the vowel for her Beijing subjects (see Figure 1).

---

11.   The appendix to Hall-Lew's (2002) study of more recent western loanwords into Mandarin drawn from the area of popular culture contains c. 130 items including the following that conform to the generalizations seen in Sections 2 and 3: *carnation* > k*ang*.*nai.xin*, *champagne* > *xi*a*ng*.*bin*, *crayon* > *gu.li.r*ong, *hamburger* > h*an*.*bao*, *nylon* > *ni.l*ong, *sandwich* > s*an*.*ming.zhi*, and *sauna* > s*ang*.*na*.
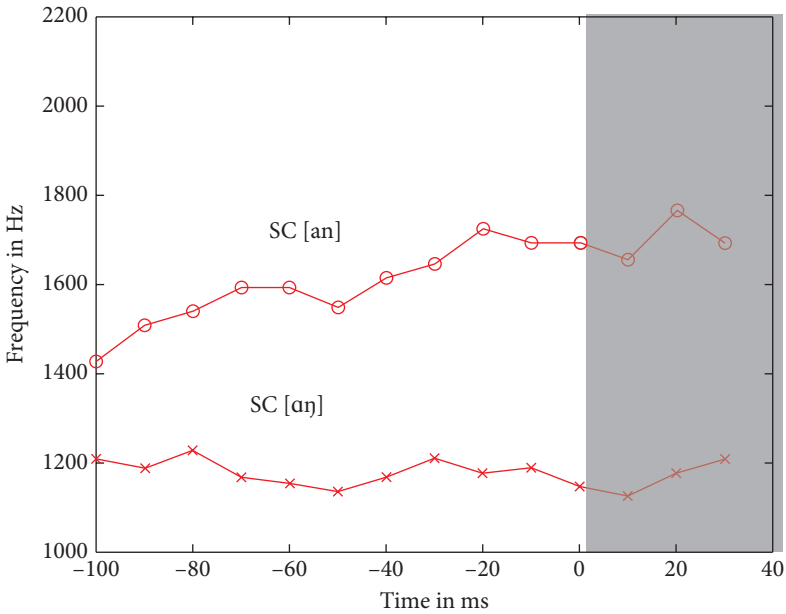
Figure 1. Averaged values of F2 movement for 18 Standard Chinese (Mandarin) vowels from 100 ms. prior to the nasal consonant to 30 ms. into the nasal consonant

Another relevant finding by Mou (2006) was that the average F2 values for pre-dorsal low and mid vowels are relatively close to the values found in syllables lacking a coda while the pre-coronal nuclei are more distant from such open syllables.

(34)   F2 in Hz
       Ca   1111   Ce   1440
       Caŋ   1172   Ceŋ   1448
       Can   1330   Cen   1578

This difference makes sense under Flemming's (2003) interpretation of the relation between coronal consonants and vocalic tongue body features as one of fronting the tongue body to accommodate a consonantal constriction at the alveolar ridge. The relatively steady rise in F2 for [an] in Figure 1 in comparsion to the largely flat trajectory in [aŋ] also makes sense in these terms. Finally, Mou (2006) reports gating experiments in which her subjects could reliably guess the presence and identity of the upcoming coda nasal when they heard less than half of a low or mid vowel. On the other hand with high vowels, where there is a contrast among [i], [y], and [u], speakers could not reliably identify the following nasal–especially after [i] where there may even be neutralization of the [n] vs. [ŋ] contrast. In sum, Rhyme Harmony is a genuine process of Mandarin grammar–an enhancement

effect (Keyser & Stevens 2006) that speakers can utilize to identify the place of articulation of the nasal coda.

5.   Summary and conclusion

This study utilized the Mandarin nasal codas to probe the phonological vs. phonetic bases of loanword adaptation. Nonhigh vowels are assigned different allophones along the front–back dimension in order to enhance a phonemic contrast between coronal and dorsal nasal codas. Our principal finding is that when the adapter is presented with conflicting choices to satisfy this phonotactic constraint of native grammar, it is the information found in the phonetically more salient vowel that determines the outcome. This result is in line with other cases of such conflict in loanword adaptation reported in Kenstowicz (2003). Coupled with the observation that stressed syllables are often the site of cyclic transfer (Kenstowicz 1997; Steriade 1999), it suggests that perceptual salience constitutes an alternative dimension of phonological faithfulness.

Tasks for future research include more extensive documentation and analysis of current loanword adaptation patterns in Mandarin as well as a more quantitative analysis of the Rhyme Harmony process along the lines of Flemming (2008). More generally, our study raises the question of whether enhancements which play a role in speech perception couple together features or cues that have a natural, cross-linguistically recurrent relation such F0 and duration cues to consonantal laryngeal contrasts (Hsieh & Kenstowicz 2008); or can they involve more phonetically arbitrary connections that are rooted in the accidents of the history of individual languages? This is of course a fundamental question that has emerged in the field of phonology more generally in the past decade. We believe that continued study of loanword adaptation may provide crucial evidence to help resolve this matter.

References

Bauer, Robert. 1985. "The expanding syllabary in Hong Kong Cantonese". *Cahiers de linguistique Asie orientale* 14.99–111.

Chen, Marilyn. 2000. "Acoustic Analysis of Simple Vowels Preceding a Nasal in Standard Chinese". *Journal of Phonetics* 28.43–67.

De Lacy, Paul. 2006. *Markedness: Reduction and Preservation in Phonology.* (=Cambridge Studies in Linguistics 112.) Cambridge: Cambridge University Press.

Duanmu, San. 2000. *The Phonology of Standard Chinese*. Oxford: Oxford University Press. (2007, 2nd edition.)

Flemming, Edward. 2003. "The Relation Between Coronal Place and Vowel Backness". *Phonology* 20.335–73.

Flemming, Edward & Sarah Johnson. 2007. "Rosa's Roses: Reduced Vowels in American English". *Journal of the International Phonetic Association* 37.83–96.

Flemming, Edward. 2008. "The Grammar of Coarticulation". To appear in *La Coarticulation: Indices, Direction et Representation*, ed. by M. Embarki & C. Dodane.

Hall-Lew, Lauren Asia. 2002. *English Loanwords in Mandarin Chinese*. BA Honors Thesis, University of Arizona, Tucson.

Hsieh, Feng-fan & Michael Kenstowicz. 2008. "Phonetic Knowledge in Tonal Adaptation: Mandarin and English Loanwords in Lhasa Tibetan". *Journal of East Asian Linguistics* 17.279–97.

Kenstowicz, Michael. 1997. "Base-identity and Uniform Exponence: Alternatives to Cyclicity". *Current Trends in Phonology: Models and Methods* ed. by Jacques Durand & Bernard Laks, vol. 1, 363–94. Salford: University of Salford.

Kenstowicz, Michael. 2003. "Salience and Similarity in Loanword Adaptation: a Case Study from Fijian". *Language Sciences* 29.316–40.

Kenstowicz, Michael. 2005. "The Phonetics and Phonology of Korean Loanword Adaptation". *Proceedings of the First European Conference on Korean Linguistics* ed. by S-J. Rhee, 17–32. Seoul: Hankook Publishing Co.

Kenstowicz, Michael & Atiwong Suchato. 2006. "Issues in Loanword Adaptation: a Case Study from Thai". *Lingua* 116.921–49.

Keyser, Jay Samuel & Kenneth Noble Stevens. 2006. "Enhancement and Overlap in the Speech Chain". *Language* 82.33–63.

LaCharité, Darlene & Carole Paradis. 2005. "Category Preservation and Proximity vs. Phonetic Approximation in Loanword Adaptation". *Linguistic Inquiry* 36.223–58.

Lin, Yen-hwei. 2008. "Variable Vowel Adaptation in Standard Mandarin Loanwords". *Journal of East Asian Linguistics* 17. 363–380.

Liu Zhengtan, Gao Mingkai, Mai Yongqian, & Shi Yowei. 1984. *Dictionary of Loanwords and Hybrid Words in Chinese*. Shanghai cishu chubanshe, Shanghai.

Mou, Xiaomin. 2006. *Nasal Codas in Standard Chinese: a Study in the Distinctive Feature Theory*. Ph.D. dissertation, Massachusetts Institute of Technology.

Novotná, Zdenka. 1967. "Linguistic Factors of the Low Adaptability of Loan-words to the Lexical System of Modern Chinese". *Monumenta Serica* 26.103–18.

Paradis, Carole & Darlene LaCharité. 1997. "Preservation and Minimality in Loanword Adaptation". *Journal of Linguistics* 33.379–430.

Silverman, Daniel. 1992. "Multiple Scansions in Loanword Phonology: Evidence from Cantonese". *Phonology* 9.289–328.

Steriade, Donca. 2000. "Paradigm Uniformity and the Phonetics-Phonology Boundary". *Acquisition and the Lexicon* ed. by Michael Broe & Janet Pierrehumbert (=Papers in Laboratory Phonology 5.), 313–34. Cambridge: Cambridge University Press.

Wang, Jenny Zhijie. 1993. *The Geometry of Segmental Features in Beijing Mandarin*. Ph.D. dissertation, University of Delaware.

Yip, Moira. 2006. "The Symbiosis Between Perception and Grammar in Loanword Phonology." *Lingua* 116.950–75.

Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation

Hyunsoon Kim
Hongik University, Seoul Korea

The present study aims to elaborate on Kim's (2007a) feature-driven model of loanword adaptation, based on Korean adaptation of English affricates and fricatives (/f, v, θ, ð, s, z, ʃ, ʒ, tʃ, dʒ/) which the host language (L1) does not possess except /s/. We propose an L1 grammar-driven perception of L2 (source language) sounds in that a Korean speakers' perception is driven by native (L1) distinctive features and syllable structure rather than in terms of the unstructured L2 acoustical input per se or of L2 phonological categories. In addition, native structural restrictions are proposed to come into play when L2 sounds scanned by L1 grammar are lexicalized as new words in L1 lexical representations. It is also suggested that an L2 acoustic signal can be constrained by L1 distinctive features by virtue of normalization or generalization, when the L2 signal has no acoustic cues to L1 distinctive features, indicating that L1 grammar exerts a force in perception.

1. Introduction

This paper is concerned with how English affricates and fricatives are borrowed into Korean, as a follow-up study to Kim's (2007a) feature-driven model of loanword adaptation. Based on the Korean adaptation of English and French voicing contrasts in plosives, Kim (2007a) has proposed that when L1 speakers perceive an L2 acoustic signal, they parse the L2 signal for cues to the distinctive features of their own phonemes. For example, as shown in (1), the English voiceless plosives /p, t, k/ are all borrowed as aspirated /pʰ, tʰ, kʰ/, no matter where they are placed in a word (a), and the French voiceless plosives are borrowed as either aspirated /pʰ, tʰ, kʰ/ or as fortis /p', t', k'/ context-freely, as in (b).[1] As for voiced /b, d, g/ in English (2a) and French (2b), they are borrowed as the Korean voiceless lenis plosives /p, t, k/.

---

1. Korean adapted forms in the present study are transcribed as lexical (underlying) representations in (4b) in a feature-driven model of loanword adaptation in (4), unless marked

(1)   a.   English words   Korean adapted forms

        type t^ha.i.p^{h}ɨ
        pentium   p^{h}ɛn.t^{h}i.ʌm
        computer   k^{h}ʌm.p^{h}u.t^{h}ʌ

    b.   French words   Korean adapted forms

        Printemps (department store)   p^{h}ɨ.lɛŋ.t^{h}aŋ   ~   p'ɨ.lɛŋ.t'aŋ
        Cannes   k^{h}an.nɨ   ~   k'an.nɨ
        boutique   pu.t^{h}i.k^{h}ɨ   ~   pu.t'i.k'ɨ

(2)   English words   Korean adapted   French words   Korean adapted
                              forms   forms

    a.   bag   **pæk**   b.   Bordeaux   **po.lɨ.to**

          dam   **tæm**   grand   **kɨ.laŋ**

          gas   **ka.s'ɨ**   De Gaulle   **tɨ.kol**

Adopting the view that Korean voiceless obstruents are specified for the laryngeal features [±tense] and [±spread glottis] (henceforth, [±s.g.]), as shown in (3), Kim (2007a) has suggested that Korean speakers scan the acoustic signal of the English and French voicing contrasts for cues to the laryngeal features of Korean consonants.[2]

(3)   The laryngeal feature specification of Korean obstruents
                         (Kim 2003, 2005 a, b; Kim et al. 2005a, b, 2007)

	lenis	fortis	aspirated
[tense]	–	+	+
[s.g.]	–	–	+

In the feature system in (3), it is proposed that the feature [tense] is defined in terms of the tensing of both the primary articulator and the vocal folds, in that closure/constriction duration and larynx raising are invariant articulatory correlates

---

or mentioned specifically. This is because it is easy for readers to see how the L2 voicing contrasts are borrowed into Korean. If we transcribed surface representations instead of underlying ones, the Korean adaptation of the L2 contrasts would be less straightforward because Korean lenis stops are optionally voiced in intervocalic position on the surface. In addition, the adaptations of relevant L2 consonants as well as L1 counterparts are marked in bold throughout the present paper for the purpose of a reader's convenience.

2.   As for other views on the feature specification of Korean consonants, see C.-W. Kim (1965), Halle & Stevens (1971), Kagaya (1974) and Y.-M. Cho (1990) among others. In the feature system in (3), Korean fortis consonants are considered as singletons, not geminates (see Cho & Inkelas 1994; Kim 2002, 2005a, b; Kim et al. 2007 for phonological and phonetic arguments contra gemination and Martin (1982); Silva (1992); Han (1996) and Avery & Idsardi (2001) for phonological arguments pro gemination).

of the feature [tense] (Kim et al. 2007) and that [s.g.] is defined in terms of glottal opening, as in Halle & Stevens (1971).[3] The main acoustic correlates of the features [tense] and [s.g.] are closure/constriction duration and aspiration, respectively.

Based on the laryngeal feature specification in (3) and the Korean adaptation of English and French voicing contrasts, Kim (2007a) has proposed that when L2 voiceless/voiced plosives have a closure duration difference with no perceived difference in aspiration, as in French, the L2 voicing contrast is interpreted as the tense vs. lax opposition ([±tense]), in terms of duration, within the framework of the L1 laryngeal features in (3). Thus, the Korean treatment of French voiceless plosives in (1b) is concerned with the feature [+tense] and that of French voiced ones in (2b) with [–tense] across all contexts. When the L2 voicing contrast is perceived as the presence vs. absence of VOT lag, as in English plosives, it is interpreted in terms of [±s.g.]. As to the closure duration difference between English voiced and voiceless plosives, it has been assumed that it reinforces a Korean speakers' perception of the English voicing contrast and that it is interpreted as cuing the feature [±tense] in Korean. Accordingly, English voiceless plosives in (1a) are adapted as aspirated ([+s.g., +tense]) and voiced ones in (2a) as lenis ([–s.g., –tense]). Note that the presence vs. absence of the vocal fold vibration in the source languages is not attended to by Korean speakers. Instead, redundant phonetic cues such as closure duration and aspiration of the source languages are readily perceived as distinctive phonetic attributes of the two features [±tense] and [±s.g.], respectively, in Korean.[4]

Given this, Kim (2007a) has proposed a feature-driven model of loanword adaptation, which is composed of three main levels between the L2 acoustic output (=L1 input) and the L1 output, as shown in (4). In this model, it is assumed that acoustic parameters and cues are extracted in the first stage of L1 perception (4a i) and that they are mapped into L1 linguistic entities such as distinctive features and syllable structure in conformity with the L1 grammar (4a ii). The bidirectional arrow is used to indicate that the extraction process is guided by the categories of the L1 grammar. In particular, it is assumed that, in the mental lexicon (4b), L2 sounds scanned by L1 grammar such as native distinctive features and syllable structure are lexicalized as new words in accordance with L1 structural restrictions and that they are represented as a sequence of syllabified distinctive feature

---

3.   Following a tradition launched by C.-W. Kim (1965), Kim et al. (2007) further elaborate articulatory correlates of the feature [tense] in Korean.

4.   See also Kim (2007b, 2008) for Korean adaptation of Japanese voicing contrast in favor of such an L1 feature-driven perception in loanword adaptation.

bundles stored in long-term memory in line with Stevens (2005).[5] The lexical representations may then undergo L1 phonology (4c) from which L1 outputs result as surface representations.

(4)   A feature-driven model of loanword adaptation (Kim 2007a)

L2 acoustic output (=L1 input)

a. **L1 Perception**

    i. extraction of acoustic parameters and cues

    ii. **L1 Grammar**
       (mapping into features and syllable structure)

b. **L1 Lexical representations**
   (mental lexicon)

c. **L1 Phonology**

L1 output
(surface representations)
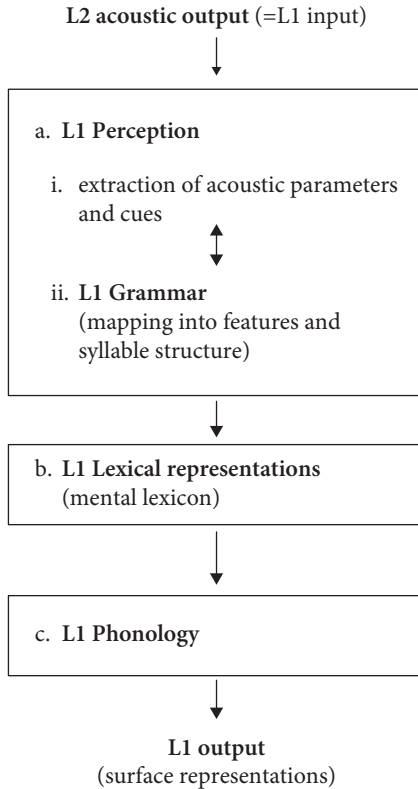
In the present study, we elaborate the feature-driven model of loanword adaptation in (4), on the basis of the Korean adaptation of English affricates and fricatives (/f, v, θ, ð, s, z, ʃ, ʒ, tʃ, dʒ/). The examination of Korean adaptation data will lead us to suggest an L1 grammar-driven perception of L2 sounds in that Korean (L1) speakers parse the acoustic signal of English (L2) affricates and fricatives within the framework of L1 distinctive features and syllable structure, rather than in terms of the unstructured L2 acoustical input per se or of L2 phonological categories. In

---

5.   Note that lexical representations in the mental lexicon (4b) are not the same as lexical and underlying representations in Lexical Phonology (e.g., Kiparsky 1982; Mohanan 1986). See the Subsection 3.4 for discussion.

addition, we suggest that L1 structural restrictions play a role in L1 lexical representation (4b) where L2 sounds scanned by L1 grammar in L1 perception are lexicalized as new words, motivating the presence of the mental lexicon (4b) between L1 perception and L1 phonology in loanword adaptation. It is also proposed that L2 acoustic signals can be constrained by virtue of normalization or generalization, when they have no acoustic cues to L1 distinctive features, indicating that L1 grammar exerts a force in perception.

This paper is structured as follows. Section 2 provides an overview of English and Korean consonants and syllable structure. Section 3 presents the Korean adaptation of English affricates and fricatives in support of the feature-driven model of loanword adaptation in (4). Section 4 discusses the phonetic approximation view and the purely phonological view of loanword adaptation in comparison with the proposal made in this study, and Section 5 is a brief conclusion.

2.   Background

The English and Korean consonant inventories are shown in (5) and (6), respectively. English has eleven affricates and fricatives (/f, v, θ, ð, s, z, ʃ, ʒ, tʃ, dʒ, h/) which have voicing contrast, like plosives, except /h/, and their places of articulation range from labial-dental (/f, v/) through denti-alveolar (/θ, ð/), alveolar (/s, z/) and palato-alveolar (/tʃ, dʒ/) to glottal (/h/).

(5)   English consonants (Ladefoged 2001)

			labial	dental	alveolar	palato-alveolar	palatal	dorsal	glottal
a.	stops	voiceless	p		t	tʃ		k	
		voiced	b		d	dʒ		g	
b.	fricatives	voiceless	f	θ	s	ʃ			h
		voiced	v	ð	z	ʒ			
c.	nasals		m		n			ŋ	
	liquid				l, r	ŋ			
	glides						j	w	

In contrast, Korean has six affricates and fricatives (/ts, tsʰ, ts’, s, s’, h/) all of which are voiceless, and the strident coronal consonants are all alveolars with a three- or two-way laryngeal contrast: lenis (/ts/), aspirated (/tsʰ/) and fortis (/ts’/) affricates and lenis (/s/) and fortis (/s’/) fricatives, as shown in (6).[6]

---

6.   The Korean affricates in (6) are transcribed as alveolar in line with Skaličková (1960) and Kim (1999, 2001, 2004) and the fricatives as lenis (/s/) and fortis (/s’/), following Kim (2005b) and Kim et al. (2005a, b) among others.

(6)   Korean consonants (Kim 2005a, b; Kim et al. 2005a, b)

		labial	coronal		dorsal	glottal
			alveolar	palatal		
a.	stops lenis	p	t, ts		k	
	aspirated	p^h	t^h, ts^h		k^h	
	fortis	p'	t', ts'		k'	
b.	fricatives lenis		s			h
	fortis		s'			
c.	nasals	m	n		ŋ	
	liquid		l			
	glides			j	w	

The syllable in Korean is composed of (C)(G)V(C) where G refers to either of the two glides /j, w/. All the consonants except /ŋ/ are allowed in onset position in Korean. Acceptable coda consonants are confined to /p, t, k, m, n, l, ŋ/ due to the Coda Neutralization process whereby all the laryngeal contrasts of obstruents are reduced to lenis counterparts.[7] For example, the fricatives /s, s', h/ and the affricates /ts, ts^h, ts'/ are neutralized into [t] in coda position like the coronal plosives /t, t^h, t'/.

The Korean adaptation data in the present study were collected from daily expressions used frequently in mass media such as advertisements in magazines, in newspapers, on the Internet or on television as well as in the spoken language.

3.   L1 grammar-driven perception of L2 sounds

In this section, we examine how the English affricates and fricatives (/f, v, θ, ð, z, ʃ, ʒ, tʃ, dʒ/) are adapted into Korean in four subsections, as follows: (a) L1 feature-driven perception of English [s], (b) the role of native (L1) syllable structure as well as distinctive features in a Korean speakers' perception of prevocalic [ʃ] and postvocalic [ʃ, ʒ, dʒ, tʃ], (c) the generalization of the feature-driven adaptation of the voicing contrast in English plosives into the adaptation of the voicing contrast in English affricates and fricatives, and (d) native structural restrictions in L1 lexical representations (4b).

3.1   L1 feature-driven perception

The Korean adaptation of English [s] supports the view of Kim (2006, 2007a,b, 2008) that the distinctive features of one's native language steer speakers in their

---

7.   See Kim & Jongman (1996) and relevant references therein for the Korean Coda Neutralization.

search to replace foreign sounds with native sounds. As shown in (7a), English single [s] is borrowed as fortis /s'/ across the contexts.[8] English [s] in consonant clusters is, on the other hand, borrowed as lenis /s/ in all contexts, as in (7b).

(7)   English words   Korean adapted forms

a.   salad   s'æl.lʌ.tɨ
       sign   s'a.in
       single   s'iŋ.kɨl
       excite   ik.s'a.ɨ.tʰɨ
       bus   pʌ.s'ɨ
       kiss   kʰi.s'ɨ

b.   sky   sɨ.kʰa.i
       snap   sɨ.næp
       disco   ti.sɨ.kʰo
       display   ti.sɨ.pʰɨl.lɛ.i
       test   tʰɛ.sɨ.tʰɨ
       mask   ma.sɨ.kʰɨ

Phonetic studies of English [s] report that the oral constriction duration is shorter in [s] in consonant clusters than in the single [s]. According to Klatt's (1976) acoustic data, the average durational reduction of English [s] in consonant clusters is "approximately 15%" compared to the single [s] in an unstressed position. In Korean the fortis fricative /s'/ is longer in constriction duration than its lenis counterpart /s/ both word-initially and -medially (e.g., Kim et al. 2005a, b), and the difference in duration between the two types of fricatives is perceived distinctively by Korean speakers in some recent perception studies. For example, in S. Kim's (1999) perception test, Koreans perceive an English longer [s] as fortis /s'/ and a shorter [s] as lenis /s/, when exposed to digitally edited [sa] stimuli. Her subjects were likely to perceive stimuli under 110 ms as the lenis fricative /s/, stimuli above 140 ms as the fortis /s'/. In addition, Lee & Iverson (2006) have reported from their perception experiments that the English fricative [s] is short enough to be perceived as /s/ when it occurs before another consonant (e.g., *stop, snap; desk, fast*) but long enough to be perceived as /s'/ when it occurs after consonants including sonorants (e.g., *dance, false, matrix*).

Given the phonetic studies and the view of the laryngeal specification of Korean obstruents in (3) as well, we propose in line with Kim (2007a) that the duration difference in English [s], which is purely phonetic in the source language,

---

8.   See Kim (2007a) for the Korean adaptation of French single [s] and [s] in consonant clusters which is the same as that of English single [s] and [s] in consonant clusters in (7).

is parsed for cues to the feature [±tense] in the initial processing of L1 perception (4a i). Accordingly, the long duration of a single [s] is parsed for a cue to [+tense] and the short one in consonant cluster to [−tense], as in (8).

(8)   English (L2)   cues   Korean (L1)

single [s]   ⟷   long duration   ⟷   [+tense]
[s] in clusters   ⟷   short duration ⟷   [−tense]

In addition, large acoustic intensity during constriction and the location of noise above 4 kHz in English [s] are extracted and parsed for cues to the Korean features [+continuant, +strident] and [+anterior, coronal], respectively, in the first stage of L1 perception (4a i). In the next stage (4a ii), where features are formally mapped in accordance with L1 grammar, the extracted cues are interpreted in terms of the L1 features. Therefore, in the lexical representations in (4b), the English [s] is represented as fortis /s'/ ([+continuant, +strident, +anterior, coronal, +tense, −s.g.]) or lenis /s/ ([+continuant, +strident, +anterior, coronal, −tense, −s.g.]), depending on whether it is long or short.[9] As a result, English [s] is realised as either fortis [s'] or lenis [s] in the Korean surface representations.

The Korean adaptation of English [s] indicates that distinctive features of the host language play a crucial role in interlanguage loanword adaptation. In the next subsection, other data is presented which shows that not only native distinctive features but also syllable structure exerts an influence on a Korean speakers' perception of English prevocalic [ʃ] and postvocalic [ʃ, dʒ, tʃ].

3.2   The role of L1 syllable structure as well as distinctive features

Prevocalic palato-alveolar [ʃ] is realised as a sequence of /s/ and /j/ before a non-front vowel, as shown in (9a) or that of /s/ and /w/ before a front vowel, as in (9b).

(9)   English words   Korean adapted forms

a.   shop   **sj**ap
     super (market)   **sj**u.pʰʌ
     special   sɨ.pʰɛ.**sj**ʌl
     audition   o.ti.**sj**ʌn
     show   **sj**o   (~ **s'j**o)
     tissue   tʰi.**sj**u   (~ tʰi.**s'j**u)
     issue   i.**sj**u   (~ i.**s'j**u)

---

9.   The Korean fricatives are specified for [−s.g.] in the laryngeal feature system in (3). See Kim et al. (2005a, b) and Kim (2005b) for both phonetic and phonological discussions for the feature specification of the Korean fricatives.

b. membership   mɛm.pʌ.**swip**
    gossip   ka.**swip**
    sheath (dress)   **swi**.tɨ
    Shell (oil company) **swɛ**l
    Sheraton (a hotel chain) **swɛ**.la.tʰon

We assume that a large acoustic intensity (stridency) of onset [ʃ] is extracted and parsed for cues to the features [+continuant, +strident] in the initial processing of L1 perception (4a i).[10] In particular, during the extraction, its low-frequency energy around 2–3.5 kHz (e.g., Kent & Read 2002) and locus are parsed for cues to the features [+anterior, coronal] and one of the glides /j/ or /w/ according to the native distinctive features and syllable structure, as shown in (10).

(10)   English (L2)   cues   Korean (L1)
              prevocalic [ʃ]   /sj/ or /sw/

     [−anterior, coronal] ⟷   noise   ⟵   [+anterior, coronal]
                              around 2–3.5 kHz   with the addition of
                              and the locus   the glide /j/ or /w/
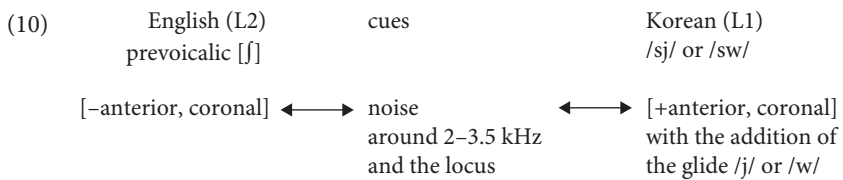
When the English fricative is followed by front vowels, the glide /j/ whose place of articulation is coronal like a front vowel is not allowed due to the violation of the Obligatory Contour Principle (henceforth, OCP. e.g., Leben 1973; McCarthy 1986; Yip 1988). Instead, the other glide /w/ is inserted by virtue of the native syllable structure, as in (9b). The extracted acoustic cues are then mapped into the L1 features and syllable structure within the framework of L1 grammar in the second stage of L1 perception (4a ii). Therefore, English [ʃ] is borrowed as either /sj/ before a non-front vowel or /sw/ before a front vowel. This results in a Korean adaptation which most closely approximates the onset palato-alveolar fricative of the source language.

     Note that length is not decisive in the adaptation of English palato-alveolar fricative in (9). The presence of the anterior lenis /s/ and fortis /s'/ in the Korean consonant inventory (6) leads to L1 feature-driven perception of English [s], as in (8), and thus Korean adapters are sensitive to the purely phonetic length difference in English [s], as in (7). However, the adaptation of English palato-alveolar fricative [ʃ] in (9) is affected by L1 syllable structure according to which one of the glides /j/

---

10. Given the feature [+anterior] is not distinctive in Korean because all coronal obstruents are [+anterior], L2 sounds which are either [+anterior] or [−anterior] are parsed for the native feature [coronal], and when they are [−anterior], the native syllable structure exerts a force, as we will see in (9) and (13).

or /w/ is inserted after the lenis /s/ to denote the place of articulation of the source sound. In addition, the onset sequences of /s'j, s'w/ are not allowed in Korean except in a few loans. For example, among the loans in (9), the Korean adaptation of [ʃ] in the English words *show, tissue* and *issue* alternates between [sj] and [s'j] on the surface. The alternation is reminicent of that between lenis and fortis consonants in intensified expressions in some native Korean words whose initial consonants are either lenis plosives or fricative /s/ in (11a) and also in some English loans (11b), as shown below. A fortis consonant in word-initial position in (11) reflects a more emphasized expression than its lenis counterpart in both some native vocabulary and English loans whose frequency is relatively higher than any others.

(11)   The alternation of the lenis and fortis consonants in native Korean words (Kim 2005a, b)

			intensified expressions		
a.	pʌn.tɛ.ki	~	p'ʌn.tɛ.ki	* pʰʌn.tɛ.ki	'chrysalis, pupa'
	tæ.tsuŋ	~	t'æ.tsuŋ	* tʰæ.tsuŋ	'Daejung (former president)'
	tsa.sik	~	ts'a.sik	*tsʰa.sik	'chap'
	koŋ.ts'a	~	k'oŋ.ts'a	*kʰoŋ.ts'a	'something got for nothing'
	so.tsu	~	s'o.tsu		'Korean distilled liquor'
	sa.raŋ	~	s'a.raŋ		'love'
	sa.na.i	~	s'a.na.i		'man'
b.	pæk	~	p'æk	*pʰæk	'bag'
	pæn.tɨ	~	p'æn.tɨ	* pʰæn.tɨ	'(musical) band'
	tæn.sɨ	~	t'æn.sɨ	* tʰæn.sɨ	'dance'
	tæm	~	t'æm	* tʰæm	'dam'
	tsæm	~	ts'æm	* tsʰæm	'jam'
	kɛ.im	~	k'ɛ.im	* kʰɛ.im	'game'
	ka.s'ɨ.	~	k'a.s'ɨ.	* kʰa.s'ɨ.	'gas'

The alternation pattern in the three English words in (9a) which is similar to that in (11) leads us to suggest that it results from a lexical diffusion of the emphasized expressions of word-initial consonants in (11). Hence, the lenis fricative of the sequence /sj/ in the English loans is emphasized into /s'j/ not only word-initially, as in [**sj**o] ~ [**s'j**o] 'show', as in (11), but also word-medially, as in [tʰi.**sj**u] ~ [tʰi.**s'j**u] 'tissue' and [i.**sj**u] ~ [i.**s'j**u] 'issue'.[11,12]

The influence of the native syllable structure as well as distinctive features on a Korean speakers'perception is further supported by the adaptation of English

---

11.   We hardly find loans beginning with /s'w/.

12.   The alternation in (11) provides evidence for the features [±tense] and [±s.g.] of the Korean obstruents (Kim 2005a, b). The sound pattern of lenis and fortis consonants to the exclusion of aspirated ones as well as the alternation is accounted for under the feature specification in (3).

words ending in palato-alveolar [ʃ, dʒ, tʃ] and anterior affricates and fricatives [s, z, f, v, θ, ð]. As shown in (12a), when the English word-final anterior affricates and fricatives are borrowed into Korean, the vowel /ɨ/ is inserted after the consonants.[13] This is also true of English word-final labial, coronal and dorsal stops, as in (12b). However, English word-final [ʃ, dʒ, tʃ] are borrowed as onsets with the insertion of the vowel /i/.[14] In particular, note the alternation between /swi/ and /si/ in the adaptation of the coda [ʃ] in (13a), and the sequence /si/ in (13a ii) is much more preferred.

(12)   The insertion of the vowel /ɨ/

       English words   Korean adapted forms

a.   rose   lo.**tsɨ**

       quiz   kʰwi.**tsɨ**

       size   sʼa.i.**tsɨ**

       bath   pa.**sʼɨ**[15]

       beef   pi.**pʰɨ**

       life   la.i.**pʰɨ**

       love   lʌ.**pɨ**

       five   pʰa.i.**pɨ**

b.   Cape (Town)   kʰɛ.i.**pʰɨ**

       print   pʰɨ.lin.**tʰɨ**

       guide   ka.i.**tɨ**

       peak   pʰi.**kʰɨ**

       gag   kæ.**kɨ**

(13)   The insertion of the vowel /i/

       English words   Korean adapted forms

                        i.   ii.

a.   cash   kʰæ.**swi**   ~   kʰæ.**si**

       rush (hour)   lʌ.**swi**   ~   lʌ.**si**

       English   iŋ.kɨl.li.**swi** ~   iŋ.kɨl.li.**si**

       Bush   pu.**swi**   ~   pu.**si**

       fish   pʰi.**swi**   ~   pʰi.**si**

---

13.   The Korean adaptation of English /z, f, v, θ, ð, ʒ, dʒ, tʃ/ will be discussed in detail in the next subsection.

14.   We have not found Korean adaptation of English words ending in [ʒ].

15.   Note that the English name *Elizabeth* is borrowed either as /ɛl.li.tsa.pɛ.sʼɨ/ or as /ɛl.li.tsa. pɛs/. When the feature bundle of /sʼ/ is syllabified as coda of a preceding syllable, the vowel /ɨ/ is not inserted, and neutralized into /t/ in L1 perception (4a). Yet, due to L1 structural restrictions in L1 lexical representations (4b), the word-final neutralized consonant is stored as the lenis fricative /s/. See the Subsection 3.4 for discussion about the word-final lenis fricative /s/ as in /ɛl.li.tsa.pɛs/.

b. change tsʰɛ.in.**tsi**
orange o.lɛn.**tsi**
cabbage kʰæ.pi.**tsi**
sponge sɨ.pʰʌn.**tsi**
page pʰɛ.i.**tsi**

c. match mæ.**tsʰi**
beach pi.**tsʰi**
pitch pʰi.**tsʰi**
coach kʰo.u.**tsʰi**
bench pɛn.**tsʰi**
catch kʰæ.**tsʰi**
touch tʰʌ.**tsʰi**

We assume that the insertion of the vowel /ɨ/ in (12) or /i/ in (13) occurs in L1 perception (4a) when the L2 word-final sounds are scanned as onsets by virtue of L1 syllable structure. The insertion of the vowel /i/ in (13) can be attributed to a Korean adapters' attempt to denote the place of articulation of the source sounds [ʃ, dʒ, tʃ] within the framework of Korean grammar (Kim 1999, 2004). Since there are only alveolar stridents except /h/ in Korean, as shown in (6), the acoustic signal of the English palato-alveolar stridents are parsed for cues to features of L1 relevant alveolar stridents /s, ts, tsʰ/ with the insertion of the vowel /i/ in the initial processing of L1 perception (4a i), as in (14).

(14)   English (L2)   cues   Korean (L1)
         [ʃ, dʒ, tʃ] in coda position

         [–anterior, coronal] ⟷   noise   ⟷   [+anterior, coronal]
                                     around 2–3.5 kHz   with the addition of /i/
                                     and the locus

Given that the vowel /i/ is similar to palato-alveolar consonants in place of articulation (e.g., Hume 1992; Clements & Hume 1995), the vowel /i/ insertion after the L1 anterior coronal consonants in (13) makes the adapted sounds as close as the L2 word-final palato-alveolar consonants. On the other hand, when English words end in anterior affricates and fricatives as well as plosives, the vowel /ɨ/ is inserted by default, as in (12). Therefore, the L2 word-final palato-alveolar sounds are perceived with the insertion of the vowel /i/, as in (13), differently from non-palato-alveolar ones in accordance with L1 syllable structure.

The preference of /si/ to /swi/ in the adaptation of English word-final [ʃ] in (13a) suggests that native distinctive features and syllable structure exert an influence on a Korean speakers' perception rather than perceptual similarity per se. The sequence of /sw/ is phonetically more faithul to the source sound than that

of /si/ in that the English fricative has a secondary articulation of labialization (e.g., Stevens 1998; Ladefoged 2001). Thus, the adaptation in (13a i) can be considered as resulting from a Korean speakers' effort to mimic the source sound as close as possible with the addition of the vowel /i/. However, the much more preferred adapted form /si/ indicates that, regardless of the labialization of the English voiceless palato-alveolar fricative, the acoustic signal of the source sound is parsed as a single /s/ with the insertion of the vowel /i/, within the framework of L1 grammar rather than perceptual similarity.[16] This is also true of the adaptation of English word-final [dʒ, tʃ] in (13b, c). Like [ʃ], the palato-alveolar affricates have a lip rounding as a secondary articulation in English. In the aspect of perceptual similarity, then, the insertion of the glide /w/ in (13b, c) would be expected. But this is not the case.

3.3 Generalization of L1 feature-driven perception of the voicing contrast in English plosives

In this subsection, we provide further support for the view that an L1 speakers' perception of the L2 acoustic output is conditioned by cues to L1 distinctive features and syllable structure. In particular, based on the Korean adaptation of the voicing contrast of English [z, ʒ, dʒ, tʃ, f, v, θ, ð], we suggest that the L2 acoustic signal can be constrained by L1 features through the normalization or generalization of the voicing contrast in English plosives, when the L2 signal has no acoustic cues to L1 distinctive features.

As shown in (15), voiced [z, ʒ, dʒ] are borrowed as lenis affricate /ts/ and voiceless [tʃ] as aspirated /tsʰ/. Note that the glide /j/ is added for the onset palato-alveolar [ʒ, dʒ, tʃ] when a following vowel is non-front, as in (9a), in order to denote the place of articulation of the stridents. No glide is inserted when the L2 sounds are followed by a front vowel, though the insertion of the glide /w/ would be perceptually closer to the source sounds.

(15) English words   Korean adapted forms

		English words	Korean adapted forms
a.	i.	zoom	**ts**um
		join	**ts**o.in
		design	ti.**ts**a.in
		pizza	pʰi.**ts**a
		music	mju.**ts**ik
	ii.	fusion	pʰju.**tsj**ʌn
		vision	pi.**tsj**ʌn

---

16. The adaptation in (13a ii) leads us to take an example of the same adaptation of the onset English /ʃ/ in /si.tʰɨ/ 'sheet'.

	iii.	manager	**mæ.ni.tsj**ʌ
		journal	**tsj**ʌ.nʌl
		junior	**tsj**u.ni.ʌ
		joy	**tsj**o.i
		jean	**ts**in
		original	o.li.**tsi**.nal
		digital	ti.**tsi**.tʰal
		General (Motors)	**ts**ɛ.nʌ.lʌl
		Jane	**ts**ɛ.in
		jam	**ts**æm
b.	i.	chocolate	**tsʰjo**.kʰol.lɛt
		chart	**tsʰja**.tʰɨ
		ketchup	kʰɛ.**tsʰjap**
		chewing (gum)	**tsʰju**.iŋ
		miniature	mi.ni.ʌ.**tsʰj**ʌ
	ii.	chip	**tsʰ**ip
		cheeze	**tsʰ**i.tsɨ
		chain	**tsʰ**ɛ.in
		chatting	**tsʰ**æ.tʰiŋ

From a phonetic point of view, it would be a little peculiar if English [tʃ] is borrowed as aspirated /tsʰ/ because the source sound has no aspiration. However, the consideration of how the voicing contrast in English plosives in (1a) and (2a) is adapted into Korean (Kim 2007a) leads us to make the following suggestion: the Korean adaptation of English voicing contrast in plosives is generalized in that of the voicing contrast in English stridents in (15), no matter whether there is no aspiration after the friction of English [tʃ]. That is, because English voiceless plosives are aspirated, as in (1a), aspiration is, now as a loan strategy, imposed on any voiceless English affricate or fricative except [s].[17]

It is noteworthy that the Korean adaptation of the voicing contrast in the English stridents in (15) is in accordance with that of the voicing contrast in English plosives: as English voiceless plosives in (1a) are adapted as aspirated plosives ([+s.g., +tense]) and voiced ones in (2a) as lenis ([−s.g., −tense]), voiceless strident [tʃ] is borrowed as aspirated ([+s.g., +tense]) /tsʰ/ and voiced [z, ʒ, dʒ] as lenis ([−s.g., −tense]) / ts/. Given this, it is assumed that Korean speakers instantiate the generalization of English voicing contrast in plosives when they perceive the stridents [z, ʒ, dʒ, tʃ] in the initial processing of L1 perception (4a i). That is, regardless of whether the English stridents have aspiration or not, the voicing contrast in the English

---

17.   Note that Korean anterior coronal fricatives /s/ and /s'/ are not distinctive in terms of aspiration (Kim 2005b; Kim et al. 2005a, b).

stridents is generalized by the feature-driven adaptation of the voicing contrast in plosives. Thus, the acoustic signal of the voiceless [tʃ] is generalized as [+s.g.] like voiceless plosives, and the voiced stridents as [−s.g.] like voiced ones, as marked by solid line arrows in (16). Moreover, the duration difference between the voiceless and voiced stridents is parsed for a cue to [±tense], enhancing the generalization of the voicing contrast, as marked by dotted line arrows in (16).[18]

(16)   English (L2)   cues   Korean (L1)

   [−voice] ([tʃ])   ◄——————►   by generalization   ◄——————►   [+s.g.]
                                              (with the enhancement of
                                              long duration)   ◄·········►   [+tense]
   [+voice] ([z,ʒ, dʒ]) ◄——————►   by generalization   ◄——————►   [−s.g.]
                                              (with the enhancement of
                                              short duration)   ◄·········►   [−tense]

In addition, a large acoustic intensity of English [z, ʒ, dʒ, tʃ] is parsed for cues to [−continuant, +strident] within the framework of the Korean feature system, regardless of whether English [z] is a fricative ([+continuant]). Among the Korean stridents in (6), it is only the lenis affricate /ts/ that is redundantly voiced in intervocalic position, just like the lenis plosives /p, t, k/. Furthermore, a high-frequency energy above 4 kHz and the locus of onset [z] is extracted and parsed for cues to [+anterior, coroanl], and a comparatively low-frequency energy around 2–3.5 kHz and the locus of onset [ʒ, dʒ, tʃ] to [+anterior, coroanl] with the addition of the glide /j/ in accordance with L1 syllable structure, when followed by a non-front vowel, as in (15a ii, iii) and (15b i). Yet, the glide /j/ is not allowed by virtue of OCP, when the stridents [ʒ, dʒ, tʃ] are followed by the vowel /i/.

   In the next processing of L1 perception (4a ii), the extracted cues in English [z] are mapped into the features [−continuant, +strident, −tense, −s.g.], whereas those in onset [ʒ, dʒ] and [tʃ] are mapped into [−continuant, +strident, −tense, −s.g.] and [−continuant, +strident, +tense, +s.g.], respectively. As a result, onset voiced [z, ʒ, dʒ] are borrowed as /ts(j)/ and voiceless [tʃ] as /tsʰ(j)/ into Korean.

   The Korean adaptation of English [f] ~ [v] contrast further supports that the generalization of English voicing contrast in plosives constrain a Korean speakers' perception of the voicing contrast of English affricates and fricatives. As shown in (17), onset labio-dental fricatives [f, v] are borrowed as aspirated /pʰ/ and lenis /p/, respectively.

---

18.   See Kim (2008) for the enhancement in loanword adaptation.

(17) English words   Korean adapted forms

a. phone   **pʰon**
    focus   **pʰo.u.kʰʌ.sʼɨ**
    after (service)   **æ.pʰɨ.tʰʌ**
    coffee   **kʰʌ.pʰi**
    soft   **sʼo.pʰɨ.tʰɨ**
    sofa   **sʼo.pʰa**
    uniform   **ju.ni.pʰom**
b. visa   **pi.tsa**
    violin   **pa.i.ol.lin**
    lavender   **la.pɛn.tʌ**
    oven   **o.pɨn**
    service   **sʼʌ.pi.sʼɨ**

We suggest that the voicing contrast in English [f] and [v] is generalized for [+s.g.] vs. [–s.g.] contrast, as marked by solid lines in (16) in the above and that the long vs. short constriction duration in [f] ~ [v] contrast is parsed for cues to [+tense] for [f] and to [–tense] for [v] as an enhancement of the voicing contrast in the initial processing of L1 perception (4a i), as marked by dotted line arrows in (16).

It is also assumed that the locus of English [f, v] (e.g., F2 transition) is parsed for cues to the feature [labial]. Given that it is only labial plosives ([–continuant, –strident]) that are specified for [labial] among Korean obstruents, the acoustic signal of less and random noise over a wide range of frequencies during constriction of the L2 sounds is constrained by the Korean feature system, being parsed for cues to [–continuant, –strident]. In the second processing of L1 perception (4a ii), the extracted cues are mapped into the features [–continuant, –strident, labial, ±tense, ±s.g.]. As a result, English [f] and [v] are borrowed as /pʰ/ and /p/, respectively, which is fit into the native feature system.

It is of interest that, though not often, the word-initial [f] in some English words can be borrowed as either /pʰ/ or /hw/, or as /hw/, as shown in (18).

(18) English words   Korean adapted forms

    fine   **pʰa.in**   ~   **hwa.in**
    family   **pʰæ.mil.li**   ~   **hwæ.mil.li**
    feminism   **pʰɛ.mi.ni.tsɨm**   ~   **hwɛ.mi.ni.tsɨm**
    fresh   **hwu.lɛ.swi**
    Fanta (soft drink)   **hwan.tʰa**
    Fiber (soft drink)   **hwa.i.pʌ**

The sequence of /hw/ would be phonetically closer to English [f] than aspirated /pʰ/ in that /h/ is a fricative ([+continuant, –strident]) like [f]. Yet, English [f] is borrowed much more as /pʰ/, as in (17a), than /hw/. From this, we may suggest that the generalization of English voicing contrast in plosives is preferred to

perceptual similarity, that is, purely phonetic approximation when [f, v] are borrowed into Korean.

On the other hand, the Korean adaptation of [θ, ð] shows that feature-driven perception is preferred to the generalization of English voicing contrast in plosives. The English voiceless [θ] is adapted as the fortis fricative /s'/, similar to a single /s/ in (7a), as shown in (19a), whereas the voiced [ð] is borrowed as lenis /t/, as in (19b).[19]

(19)	English words	Korean adapted forms
a.	theraphy	s'ɛ.lʌ.pʰi
	three	s'ɨ.li
	think	s'iŋ.kʰɨ
	Anthony	æn.s'o.ni
	something	s'ʌm.s'iŋ
	thank you	s'æŋ.kʰju
	bath	pa.s'ɨ
	health (club)	hɛl.s'ɨ
b.	this	**ti.s'ɨ**
	the (Body Shop)	**t**ʌ
	Brother (brand name)	pɨ.la.**t**ʌ
	smoothie	sɨ.mu.**t**i
	smooth	sɨ.mu.**t**ɨ

We assume that the presence of acoustic intensity in voiceless [θ] is parsed for cues to the features [+continuant, +strident], and its long constriction duration to a cue to the feature [+tense] like /s'/, as in (8), in accordance with the Korean feature system. A higher second formant of a neighboring vowel after the [θ, ð] than the labio-dental fricatives [f, v] would lead Korean speakers to parse it to cues to the feature [coronal]. As a result, the voiceless [θ] is borrowed as /s'/ into Korean. In the case of the English [ð], the near absence of acoustic intensity during oral constriction is parsed for cues to the features [–continuant, –strident], its short duration than [θ] to [–tense]. Thus, English [ð] is borrowed as lenis /t/ into Korean.

3.4 Native structural restrictions in L1 lexical representations

In this subsection, we examine Korean adaptation of English word-final coronal plosive consonants and propose that native structural restrictions play a role in L1 lexical representations (4b) where L2 sounds scanned by L1 grammar in L1 perception are lexicalized as new words.

---

19. There are a few exceptions for this. The English expression *thank you* is sometimes borrowed as [t'æŋ.kʰju], and the word *smooth* as [sɨ.mu.**s'**ɨ].

When English words end in non-coronal consonants [p, k, g], the consonants are borrowed either as /pʰ/, /kʰ/ and /k/, respectively, in onset position with the vowel /ɨ/ insertion, as shown in (20a i), or as /p/ and /k/ in coda position with no vowel insertion, as in (20a ii). When followed by a vowel-initial suffix such as the subject marker /-i/, the object marker /-ɨl/ and the locative marker /-ɛ/ in L1 phonology (4c), the word-final coda consonants /p, k/ in (20a ii) are syllabified as onset consonants, as in (20a iii). In the case of English word-final coronal consonants [t] and [d], they are borrowed either as /tʰ/ and /t/, respectively, in onset position with the the /ɨ/ vowel insertion, as shown in (20b i). Yet, the coronal consonants are borrowed as the lenis fricative /s/, not /tʰ/ or /t/, in coda position with no vowel insertion, as in (20b ii). Therefore, when followed by the same vowel-initial suffixes in L1 phonology (4c), the fricative /s/ is syllabified as onset consonants, as in (20a iii). It is noteworthy that the word-final consonants in (20b ii) surface as [s], not [s'], different from those in [pʌ.s'ɨ] 'bus' and [kʰi.s'ɨ] 'kiss', as in (7a).

(20)   English words   Korean adapted forms

		i.		ii.	iii. L1 phonology		
a.	soup	su.pʰɨ	~	sup	su.pi	su.pɨl	su.pɛ
	tip			tʰip	tʰi.pi	tʰi.pɨl	tʰi.pɛ
	kick			kʰik	kʰi.ki	kʰi.kɨl	kʰi.kɛ
	rock (music)			lok	lo.ki	lo.kɨl	lo.kɛ
	tag	tʰæ.kɨ	~	tʰæk	tʰæ.ki	tʰæ.kɨl	tʰæ.kɛ
	(hot)dog	to.kɨ	~	tok	to.ki	to.kɨl	to.kɛ
b.	cut	kʰʌ.tʰɨ	~	kʰʌs	kʰʌ.si	kʰʌ.sɨl	kʰʌ.sɛ
	set	s'ɛ.tʰɨ	~	s'ɛs	s'ɛ.si	s'ɛ.sɨl	s'ɛ.sɛ
	shot	sja.tʰɨ	~	sjas	sja.si	sja.sɨl	sja.sɛ
	robot	lo.po.tʰɨ	~	lo.pos	lo.po.si	lo.po.sɨl	lo.po.sɛ
	(deep) throat	s'ɨ.lo.tʰɨ	~	s'ɨ.los	s'ɨ.lo.si	s'ɨ.lo.sɨl	s'ɨ.lo.sɛ
	(i-) pod	pʰa.tɨ	~	pʰas	pʰa.si	pʰa.sɨl	pʰa.sɛ

With respect to the adaptation of English word-final plosives in onset position with the vowel /ɨ/ insertion in (20 i), we assume that it is L1 perception (4a) that the plosives which are perceived as aspirated for English voiceless ones and as lenis for voiced ones, as in (1a) and (2a), respectively, are syllabified as onset with the vowel insertion and then that the L2 sounds are stored as new words in L1 lexical representations, as in (20 i).[20] As for English loans in (20 ii), we assume that it is due to a Korean adapters' effort to mimic English word-final plosives, as closely as

---

20.   See Kang (2003) for a phonetic view on the presence/absence of the vowel /ɨ/ insertion in Korean adaptation of English words and Kim (2007a) for a different interpretation.

possible such that English word-final plosives are linked to coda position with no vowel insertion. Then the English non-coronal plosives which are borrowed as /pʰ/ and /kʰ/ are subject to the L1 phonological process of Coda Neutralization in L1 perception (4a).[21] Therefore, the loans with the neutralized word-final plosives /p/ and /k/ are lexicalized in L1 lexical representations as new words, as in (20a ii).

On the other hand, the adaptation of English word-final coronal plosives as the lenis fricative /s/ in (20b ii) is attributed to the effect of L1 structural restrictions in L1 lexical representations. That is, when linked to coda position in L1 perception, the /tʰ/ and /t/ which Korean adapters perceive for English word-final coronal plosives [t] and [d], respectively, are subject to Coda Neutralization, being reduced to /t/. But the neutralized coronal consonant is not allowed in L1 lexical repesentations (4b), due to the native structural restriction that a Korean word is likely to end with the lenis fricative /s/ rather than with the coronal plosives /t, tʰ, t'/ in the lexicon.[22,23] Thus, the English loans in (20b ii) have the lenis fricative /s/ in word-final position in L1 lexical representations. In L1 phonology, then, the lenis fricative /s/ is syllabified as an onset, as in (20b iii), when followed by a suffix beginning with a vowel. Since English word-final coronal plosives are lexically stored as /s/ in L1 lexical representations, as in (20b ii), they do surface as [s], not [s'] in onset position when followed by a vowel-initial suffix, as in (20b iii). However, when no suffix follows it in L1 phonology, the word-final /s/ in (20b ii) would undergo the L1 phonological process of Coda Neutralization, as shown in (21 ii), and surface as [t].

(21) i. L1 lexical representations ii. L1 phonology

	i. L1 lexical representations	ii. L1 phonology
cut	kʰʌs	kʰʌt
set	s'ɛs	s'ɛt
shot	sjas	sjat
robot	lo.pos	lo.pot
(deep) throat	s'ɨ.los	s'ɨ.lot
(i-)pod	pʰas	pʰat

---

21. Alternatively one might suggest that the L1 phonological process of Coda Neutralization applies only in L1 phonology (4c). This alternative, however, would be hard to account for why the lenis consonants /p, k/ come into onset position when followed by a vowel-initial suffix in (20a iii).

22. The structural restriction is caused by native-word frequency. Though there are a few native words ending in the coronal plosive /t/ or /tʰ/ (e.g., /kot/ 'immediately', /tat-/ 'to close', /tot-/ 'to rise', /sotʰ/ 'pot', /katʰ-/ 'to be the same'), most of Korean words end in /s/ in the lexicon. No Korean words end in /t'/.

23. See also Boersma & Hamann in this volume for the role of native structural restrictions in loanword adaptation.

The presence of the lenis fricative /s/ (20b ii) for English word-final coronal plosives suggests that there is an intermediate level of L1 lexical representations between L1 perception (4a) and L1 phonology (4c) in loanword adaptation, as shown in (4). If we assume that the L2 acoustic signal is scanned by L1 grammar and then subject to L1 phonology without the level of L1 lexical representations, the alternation between /s/ (20b iii) and /t/ (21 ii) in L1 phonology cannot be accounted for. Moreover, the adaptation data in (20b) suggest that L1 lexical representations (4b) are different from lexical and underlying representations in Lexical Phonology (e.g., Kiparsky 1982; Mohanan 1986). In the view of Lexical Phonology, the derivation of a word involves a shuttling back and forth between the morphological component and the phonological component through several levels, and each input to the phonology at each level is considered as an underlying form. In L1 lexical representations (4b), however, L2 sounds scanned by L1 grammar are stored as new words in accordance with L1 structural restrictions and their representations as a sequence of syllabified distinctive feature bundles are lexical or underlying forms which may undergo L1 phonology (4c). For example, the lexical representations of the loans with the lenis fricative /s/ in word-final position in (20b ii) are underlying representations which undergo L1 phonology, as in (20b iii) and (21ii).

So far we have examined how English affricates and fricatives [f, v, θ, ð, s, z, ʃ, ʒ, tʃ, dʒ] are borrowed into Korean in Kim's (2007a) feature-driven model of loanword adaptation. We have proposed that Korean speakers parse the acoustic signal of the source sounds within the framework of the native distinctive features and syllable structure, rather than in terms of the unstructured L2 acoustical input per se or of L2 phonological categories. We have also proposed that the feature-driven perception of the voicing contrast in English plosives is generalized in a Korean speakers' perception of the voicing contrast in English [f, v, z, ʃ, ʒ, tʃ, dʒ]. Furthermore, native structural restrictions are proposed to come into play in L1 lexical representations, motivating the presence of L1 lexical representations between L1 perception and L1 phonology in the model of loanword adaptation in (4).

4. Discussion

In this section, we discuss the phonetic approximation and the purely phonological views of loanword adaptation that are found in the literature in comparison with the analysis proposed in the present study and suggest that the present view in this study supports an intermediate view of loanword adaptation, that is, L1 grammar-driven perception of L2 sounds.

In the phonetic approximation view (e.g., Silverman 1992; Yip 1993; Kenstowicz 2003; Peperkamp & Dupoux 2003; Hsieh et al. in this volume), when confronted

with an L2 segment whose feature matrix does not exist in L1, L1 speakers will perceive and produce the native segment which most closely approximates the input in articulatory and/or acoustic properties. For example, we have noted that coda English [ʃ] is borrowed as either /swi/ or /si/, as in (13a). In the phonetic approximation view, the sequence [swi] would be the best candidate for the source sound in Korean adaptation. However, this is not always the case. The sequence /si/ is much more preferred. This is remenicent of the Korean adaptation of English [f] as either /pʰ/ or /hw/, as shown in (17) and (18). While /hw/ would be perceptually more similar to the source sound, /pʰ/ is more preferred. In addition, we have noted that in the Korean adaptation of the voicing contrast in English stridents [z, ʒ, dʒ, tʃ] in (15), the English voiceless strident is borrowed as /tsʰ/, though the source sound has no aspiration. This indicates that the adaptation is not based on phonetic or perceptual similarity to L2 sounds. Rather we have suggested that the feature-driven adaptation of the voicing contrast in English plosives is generalized in the adaptation of the voicing contrast in the English affricates and fricatives. This shows that L1 grammar, that is, L1 distinctive features play a crucial role in the adaptation.

In the purely phonological view, loanword adaptation is based on phonological category mappings between L2 and L1 (e.g., Paradis & LaCharité 1997; LaCharité & Paradis 2005; Paradis & Tremblay in this volume). In this view, Korean adapters are expected to make phonological category mappings between English and Korean consonants. Thus, we could expect English [s] to be borrowed as /s/ into Korean, because the two sounds are categorically the same in that it is an anterior coronal strident fricative in the two languages. However, the Korean adaptation of the source sound as /s/ or /s'/ in (7) makes it evident that what is concerned here is not phonological mappings but phonetic difference in duration which is not categorical or phonemic in English. Hence, as shown in (8), the duration difference of English [s] is parsed for cues to the Korean feature [tense].

Moreover, the examination of how plural forms of English words are borrowed into Korean reveals that it is the L2 acoustic signal, not phonological representation, that is parsed within the framework of Korean grammar. As shown in (22), the sequence of the English stem-final [t] and the plural suffix [s] is borrowed as aspirated affricate /tsʰ/ with the insertion of the vowel /ɨ/.

(22) English words   Korean adapted forms

a.   results   li.tsʌl.**tsʰɨ**
      fruits   pʰu.lu.**tsʰɨ**
      (off) limits   li.mi.**tsʰɨ**
      pants   pʰæn.**tsʰɨ**
      sports   sɨ.pʰo.**tsʰɨ**
b.   Cats (musical)   kʰæ(**t**).**tsʰɨ**
      nuts   nʌ(**t**).**tsʰɨ**

If phonological category mappings between L2 and L1 are concerned in loanword adaptation, it would be hard to account for why the English stem-final consonant and the plural suffix are borrowed into Korean as the sequence of the aspirated affricate /tsʰ/ and the vowel /ɨ/. The data in (22) indicate that the morphological information of the source language is not attended to by Korean speakers. Rather the sequence of the source sounds is perceived as the single anterior coronal aspirated affricate, and the vowel /ɨ/, not /i/ as in (13c), is inserted after the anterior affricate to meet L1 syllable structure, as in (12a). In addition, it is noteworthy that the coda consonant /t/ optionally precedes the adapted sequence of the aspirated affricate /tsʰ/ and the vowel /ɨ/, as shown in (22b). The presence of the coda consonant /t/ in the adaptation is neither phonetically nor phonologically faithful to the source sounds, because there is only the word-final coronal plosive [t] in the L2 acoustic signal of the English words [kʰæts] 'cats' and [nʌts] 'nuts' as well as in their phonological representations.

The adaptation in (22) follows if we assume that loanword operations proceed from L2 phonetic outputs, and not L2 phonological categories and that L2 phonetic outputs are constrained by L1 distinctive features and syllable structure. That is, the acoustic signal of English word-final [t] and plural suffix [s] is parsed for cues to the Korean features [−continuant, +strident, +tense, +s.g.] in the initial processing of L1 perception (4a i). In particular, the generalization of English voicing contrast holds on, such that the acoustic signal of the voiceless consonants [t] and [s] is constrained by L1 distinctive features [+tense, +s.g.]. The vowel /ɨ/ is inserted by default, as in (12), to preserve the adapted sound in onset position by virtue of L1 syllable structure. In the second processing (4a ii), where features and syllable structure are formally mapped in accordance with L1 grammar, the English sequence is mapped into the features [−continuant, +strident, +tense, +s.g.]. As a result, the sequence of the English stem-final [t] and the plural suffix [s] is borrowed as /tsʰɨ/, as in (22).

With respect to the coda consonant /t/ in (22b), we suggest that it results from the influence of L1 syllable structure on an L1 speakers' perception of L2 sounds. According to the place markedness constraint in coda position in Korean, the coronal plosive is unmarked, with the dorsal one being more marked than the labial one (e.g., Cho 1990; Jun 1995), and the unmarked coronal plosive is often deleted when followed by a consonant in Korean (e.g., Kim-Renaud 1974). We assume that the L1 syllable structure constraint in coda position also affects a Korean speakers' perception of the English words in (22b) in L1 perception (4a).

---

24. See also Kim (2008) for a similar adaptation of Japanese geminates.

Thus, the coronal plosive is in coda position or deleted within the framework of Korean syllable structure.[24]

So assuming that loanword phonology proceeds from L2 acoustic signal possessing, not from phonological structure and that the generalization of the voicing contrast in English plosives is instantiated according to the system of L1 distinctive features, we can account for the Korean adaptation in (22) in an intuitive manner. The influence of L1 features/syllable structure on an L1 speakers' perception of L2 acoustic signals in the present study supports an L1 grammar-driven perception of L2 sounds (e.g., Kim 2006, 2007a,b, 2008), in support of Polivanov (1931), Trubetzkoy (1939) and Hyman (1970) among others.[25]

5.   Conclusion

In the present study, we have looked into how English affricates and fricatives are borrowed into Korean. Based on the adaptation data, we have proposed that L2 acoustic signals are parsed for cues to L1 distinctive features, that not only L1 distinctive features but also syllable structure plays a cruical role in a Korean speakers' perception and that the feature-driven adaptation of the voicing contrast in English plosives is generalized in the adaptation of the voicing contrast in English affricates and fricatives, when they have no acoustic cues to L1 distinctive features. We have also proposed that L1 structural restrictions come into play in L1 lexical representations where L2 sounds scanned by L1 grammar are lexicalized as new words.

Some theoretical implications can be drawn. First, the present study confirms the claim that distinctive features play an explicit and crucial role in interlanguage loanword adaptation, as in Kim (2006, 2007a,b, 2008). Second, our view that an L1 speakers' perception of L2 sounds is made within the framework of L1 grammar, that is, L1 distinctive features and syllable structure supports a traditional insight going back to Polivanov (1931), Trubetzkoy (1939) and Hyman (1970) among others.

---

25.   See also Kenstowicz (2005), Kubozono (2006) and Yip (2006).

References

Avery, P. & W.J. Idsardi. 2001. "Laryngeal dimensions, completion and enhancement". *Phonology and Phonetics* ed. by T. Hall, 41–70. Berlin & New York: Mouton de Gruyter.

Boersma, P. & S. Hamann. 2008. "Loanword adaptation as first-language phonological perception". To appear in this volume.

Cho, Y.-M. 1990. *Parameters of consonantal assimilation*. Ph.D. dissertation, Stanford University.

Cho, Y.-M. & S. Inkelas. 1994. "Post-obstruent tensification in Korean and geminate inalterability". *Theoretical Issues in Korean linguistics* ed. by Y.-K. Kim-Renaud, 45–60. Stanford University: CSLI.

Clements, G.N. & E. Hume. 1995. The internal organization of speech sounds. *The Handbook of Phonological theory* ed. by John A. Goldsmith, 245–306. Oxford: Blackwell.

Davis, S. & M.-H. Cho. 2006. "Phonetics versus phonology: English word final /s/ in Korean loanword phonology". *Lingua* 116. 1008–1023.

Halle, M. & K.N. Stevens. 1971. "A note on laryngeal features". *Quarterly progress report* 101. 198–212. Cambridge, Mass.: Research Laboratory of Electronics, MIT.

Han, J.-M. 1996. *The Phonetics and Phonology of Tense and Plain Consonants in Korean*. Ph.D. dissertation, Cornell University.

Hume, E. 1992. *Front vowels, coronal consonants and their interaction in nonlinear phonology*. Ph.D. dissertation, Cornell University.

Hyman, L. 1970. "The role of borrowing in the justification of phonological grammars". *Studies in African Linguistics* 1. 1–48.

Iverson, G. & A. Lee. 2006. "Perception of contrast in Korean loanword adaptation". *Korean Linguistics* 13. 49–87.

Jun, J. 1995. *Perceptual and Articulatory Factors in Place Assimilation: an Optimality Theoretic approach*. Ph.D. Disseration, UCLA.

Kagaya, R. 1974. "A fiberscopic and acoustic study of the Korean stops, affricates and fricatives". *Journal of Phonetics* 2. 161–180.

Kang, Y. 2003. "Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean". *Phonology* 20. 219–273.

Kenstowicz, M. 2005. "The phonetics and phonology of Korean loanword adaptation". *Proceedings of the First European Conference on Korean linguistics*, ed. by S.-J. Rhee, 17–32. Seoul: Hankook Publishing co.

Kent, Ray D. & C. Read. 2002. *Acoustic Analysis of Speech*. San Diego, California: Singular Publishing Group, Inc. 2nd edition.

Kim, C.-W. 1965. "On the autonomy of the tensity feature in stop classification". Word 21. 339–359.

Kim, H. 1999. "The place of articulation of Korean affricates revisited". *Journal of East Asian Linguistics* 8. 313–347.

———2001. "A Phonetically-based Account of Phonological Assibilation". *Phonology* 18. 81–108.

———2002. "Korean tense consonants as singletons". *Proceedings of the 38th Conference of Chicago Linguistic Society*, 329–344. Chicago: Chicago University Press.

———2003. "The feature [tense] revisited: the case of Korean consonants". *Proceedings of the 34th annual meeting of the Northeast Linguistic Society*, 319–332.

———2004. "Stroboscopic-cine MRI data on Korean coronal plosives and affricates: implications for their place of articulation as alveolar". *Phonetica* 61. 234–251.

———2005a. "The representation of the three-way laryngeal contrast in Korean consonants". *The internal organization of phonological segments*, eds. by Marc van Oostendorp & Jeroen van de Weijer, 263–293. Berlin & New York: Mouton de Gruyter.

———2005b. "What features underline the /s/ vs. /sʼ/ contrast in Korean? Phonetic and Phonological evidence". To appear in *Where do features come from? The nature and sources of phonological primitives,* eds. by G.N. Clements & R. Ridouane.

——— 2006. "Feature-driven non-native perception in loanword adaptation of the English and French voicing contrast into Korean". Paper presented at the 14th Manchester Phonology Meeting in May 2006.

———2007a. "A feature-driven model of loanword adaptation: evidence from Korean". ms., Hongik University (submitted).

———2007b. "Feature-driven non-native perception in the loanword adaptation between Japanese and Korean". *Phonological Studies* 10 ed. by the Phonological Society of Japan, 77–88. Tokyo: Kaitakusha.

———2008. "Loanword adaptation between Japanese and Korean: Evidence for L1 feature-driven perception". *Journal of East Asian Linguistics* 17. 331–346.

———& A. Jongman. 1996. "Acoustic and Perceptual Evidence for Complete Neutralization of Manner of Articulation in Korean". *Journal of Phonetics* 24. 295–312.

———S. Maeda & K. Honda. 2005a. "The laryngeal characterization of Korean fricatives: stroboscopic cine-MRI data" (submitted).

———S. Maeda, K. Honda & S. Hans. 2005b. "The laryngeal characterization of Korean fricatives: acoustic and aerodynamic data". To appear in *Interdisciplinary approaches to turbulent sounds*, ed. by Fuchs, S., Toda, M. & M. Zygis. Berlin & New York: Mouton de Gruyter.

———S. Maeda & K. Honda. 2007. "Invariant articulatory bases of the features [tense] and [spread glottis] in Korean: new stroboscopic cine-MRI data". To appear in *Journal of Phonetics*.

Kim-Renaud, Y.-K. 1974. *Korean Consonantal Phonology*. Ph.D. Dissertation, University of Hawaii.

Kim, S. (1999). *Subphonemic duration difference in English /s/ and few-to-many borrowing from English to Korean*. Ph.D. dissertation, University of Washington.

Kiparsky, P. 1982. "Lexical morphology and phonology". *Linguistics in the morning calm,* vol. 2 ed. by I.-S. Yang, 3–91. Seoul: Hanshin.

Klatt, D.H. (1976). "Vowel lengthening is syntactically determined in a discourse". *Journal of Phonetics* 3. 129–140.

Kubozono, H. 2006. "Where does loanword prosody come from?: A case study of Japanese loanword accent". *Lingua* 116. 1140–1170.

LaCharité, D. & C. Paradis. 2005. "Category Preservation and Proximity versus Phonetic Approximation in Loanword Adaptation". *Linguistic Inquiry* 36. 223–258.

Ladefoged, P. 2001. *A course in phonetics.* Hartcourt College Publishers. 4th edition.

Leben, W. 1973. *Suprasegmental Phonology*. Ph.D. dissertation, MIT, Cambridge, MA.

Lee, A. & G.K. Iverson. 2006. Variation in Korean integration of English word-final /s/. *Language Research* 42. 239–251.

McCarthy, J. 1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* 17. 207–263.

Martin, S.E. 1982. "Features, Markedness, and Order in Korean Phonology". *Linguistics in the Morning Calm*, ed. by I. Yang, 601–618. Seoul: Hanshin.

Mohanan, K.P. 1986. *The Theory of Lexical Phonology*. Dordrecht: D. Reidel Publishing co.

Paradis, C. & D. LaCharité. 1997. "Preservation and minimality in loanword adaptation". *Journal of Linguistics* 33. 379–430.

Peperkamp, S. & E. Dupoux. 2003. "Reinterpreting loanword adaptations: The role of perception". In *Proceedings of the 15th International Congress of Phonetic Sciences*, ed. by M.J. Solé, D. Recasens & J. Romero, 367–370. Barcelona.

Polivanov, E. 1931. La perception des sons d'une langue étrangère [The perception of non-native language sounds]. *Travaux du Cercle Linguistique de Prague* 4. 79–96.

Skaličková, A. 1960. *The Korean consonants*. Prague, CSAV, Rozpray.

Silva, D. 1992. *The Phonetics and Phonology of Stop Lenition in Korean.* Ph.D. dissertation, Cornell University.

Silverman, D. 1992. "Multiple scansions in loanword phonology: evidence from Cantonese". *Phonology* 9. 289–328.

Stevens, K.N. 1998. *Acoustic Phonetics*. Cambridge, Mass.: MIT Press.

Stevens, K.N. 2005. "Features in speech perception and lexical access". *The Handbook of Speech Perception*, eds. by David B. Pisoni & Robert E. Remez, 125–155. Oxford: Blackwell.

Trubetzkoy, N.S. 1939. *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht. [English edition: *Principles of Phonology*. tr. C.A.M. Baltaxe. Berkeley and Los Angeles: University of California, 1969.]

Yip, M. 1988. The Obligatory Contour Principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19. 65–100.

———1993. "Cantonese loanword phonology and Optimality Theory". *Journal of East Asian Linguistics* 2. 261–291.

———2006. "The symbiosis between perception and grammar in loanword phonology". *Lingua* 116. 950–975.

The role of underlying representations in L2 Brazilian English

Andrew Nevins & David Braun
Harvard University

1. Overview

In this paper we examine two phenomena in the phonology of the English spoken by native speakers of Brazilian Portuguese (henceforth, Brazilian Portuguese English: BPE). The first is the phenomenon of "spurious affrication" before [u], a phenomenon in which English sequences of [tu] are rendered [tʃu] in BPE as well as in English loanwords adapted into Brazilian Portuguese (henceforth BP). In discussing this phenomenon we provide additional backgroud and exemplification of affrication in BP and BPE. The second phenomenon we discuss is "rhotic hypercorrection", a phenomenon in which English word-initial [h] is rendered in BPE as [r]. Spurious affrication and rhotic hypercorrection present a problem for models of loanword phonology such as LaCharité and Paradis (2005), which propose that speakers map the L2 input onto the closest phonological analogue in their L1, since BP contains the sequences [tu] and initial [h] in its native phonology. These phenomena are especially interesting in light of the fact that both phonetic approximation and the orthography of English would militate against spurious affrication and rhotic hypercorrection. We propose an explanation in terms of the underlying representations (URs) that BPE speakers adopt, thereby highlighting a crucial role for URs in L2 and loanword phonology.

The data discussed here come from a variety of sources, and reflect both systematic and sporadic occurrences, observed both in Boston's large Brazilian immigrant community and in various cities in Brazil. The phenomena described here may occur both in formal and casual speech registers. The phenomena of spurious affrication and rhotic hypercorrection were noticed independently by both authors and the salience of these observations, heretofore unexplored in the literature, provided the inspiration for the present paper. We have noticed no particular demographic variables that correlate with the occurrence of these phenomena. Due to the composition of the Boston Brazilian community, most of the observations reported here were produced by speakers originally from the state of Minas Gerais.

2.   Spurious affrication

In this section we report on the existence of BPE productions in which the sequence *tu* in English surfaces with an affricated coronal consonant instead. To understand why this puzzling and BPE-specific L2 phenomenon occurs, it is necessary to introduce first the distribution of affrication within the native phonology.

2.1   Background on BP affrication

BP has the seven underlying oral vowels in (1), as well as nasalized counterparts of the cardinal five.

(1)   i u
        e o
        ɛ ɔ
        a

The rule of affrication in BP is much like that of other languages:

(2)   Affrication: /t,d/→[tʃ,dʒ] / _[−cons, +high, −back]

The effects of affrication can be seen in the following examples. In the transcriptions that follow, we include the effects of vowel reduction, raising a final unstressed /o/ to [u] and /e/ to [i]. Note that vowel reduction feeds affrication.

(3)   a.   *ticket* [tʃi.ke.tʃi] 'ticket'
        b.   *tirar* [tʃi.ɾax] 'to take out'
        c.   *mente* [mẽ.tʃi] 'mind'
        d.   *sede* [se.dʒi] 'thirst'
        e.   *Diogo* [dʒi.o.gu] (Proper name)
        f.   *mastigar* [mas.tʃi.gax] 'to chew'
        g.   *cestinha* [ses.tʃĩ.ɲa] 'basket (dim.)'

Affrication in BP is described in Cagliari (1997). For discussion of the phonology and phonetics of affrication more generally, see Calabrese (2005), Hall and Hamann (2006), and Kim (2001). Crucially, *affrication in BP does not apply before the vowel [u]*.

(4)   a.   *tucano* [tu.kẽ.nu] 'toucan'
        b.   *turma* [tuɣ.ma] 'group'
        c.   *costurar* [kos.tu.ɾax] 'to sew'
        d.   *mistura* [mis.tu.ɾa] 'mixture'
        e.   *gato* [ga.tu] 'cat'

2.2   Spurious affrication in BPE

The curious phenomenon of BPE spurious affrication is the following: BPE speakers produce English words with the sequence /tu/ as [tʃu], *even though affrication does not apply before* /u/ *in BP*:

(5)   Spurious affrication of English [tu]-sequences:
   a.   *to, two* [tʃu]
   b.   *U2* [ju.tʃu] or [iw.tʃu]
   c.   *student* [stʃu.dənt] or [is.tʃu.dənt]
   d.   *stew* [stʃu] or [is.tʃu]
   e.   *stupid* [stʃu.pɪd] or [is.tʃu.pɪd]
   f.   *during* [dʒu.rɪŋ]

Our proposal is that spurious affrication results from an underlying representation in which these sequences are not, in fact, represented as /tu/. It is a fact that the English [u] is far more fronted than the BP [u], particularly after coronal consonants. It is probably more appropriate to transcribe the English back rounded vowel as [tu̜] (or [tʉ]). How does a BPE speaker deal with an incoming token of English [tu̜] (or [tʉ])? Contra the predictions of LaCharité and Paradis (2005), they are not simply matching it with the BP [u]. This is a pervasive loanword phenomenon. Any radio announcer or schoolkid in Brazil will pronounce the name of the band U2 as [ju.tʃu], with an affricate. BPE speakers are attempting to approximate the phonetic realization of [tu] (or [tʉ]). However, our proposal is that they must do so *using the resources available to them in their native language* (see also Boersma & Hamann, this volume). Since BP does not contain [tu̜] (or [tʉ]), those are not an option for the underlying form of *two*. We propose that BPE speakers approximate the fronted quality of English [tu̜] (or [tʉ]) by setting up an underlying representation with a non-nuclear [i].[1] The underlying forms of the words in (5) are thus the following:

(6)   BPE speakers' URs for English [tu]-sequences:
   a.   *to, two, too* /ti̯u/
   b.   *U2* /i̯u.ti̯u/
   c.   *student* /sti̯u.dənt/
   d.   *stew* /sti̯u/
   e.   *stupid* /sti̯u.pɪd/
   f.   *during* /di̯u.rɪŋ/

---

1.   Some of our consultants, when asked why they produce spurious affrication, tell us that they hear an *i-zinho* 'a little *i*' in English words like *two*.

In fact, the use of /i̯u/ to approximate post-coronal allophones of English [ʉ] in the underlying representations of BPE is not unique to the stops. BPE speakers adopt /i̯u/ for many post-coronal occurrences of English /u/. It is only with the stops that the non-nuclear high front vocoid /iu/ triggers affrication. With other preceding coronals, the /i̯u/ remains, and many speakers enact a process of "fusion" whereby the [−back, +high] non-nuclear segment and the [+round, +high] nucleus fuse into a front rounded vowel:

(7)   a.   *new* /ni̯u/ → [nü]
      b.   *soon* /si̯un/ → [sün]
      c.   *noon* /ni̯un/ → [nün]

The fact that the sequence /i̯u/ triggers affrication in the case of a preceding stop but triggers nuclear fusion in the case of a preceding nasal or fricative suggests a context-sensitive resolution of the same underlying sequence in different ways, suggesting that loanwords are being adopted according to an active model of "analysis-by-synthesis", as proposed by Calabrese (this volume).

A word is in order here about other possible explanations for spurious affrication. One might think that spurious affrication in words like *two* (English [tʰu̯]) represents an attempt to approximate the aspiration of the voiceless stop by using the resources of the native language, i.e. the turbulence of an affricate. While interesting, this possibility stops short when it comes to *st* words like *student*, which are not aspirated in English. It also could not be extended to the cases of affrication with voiced coronal stops, such as *during*.[2] Moreover, we dismiss the possibility that representations such as /tu̯/ come from hearing dialects of English that allow [tju] in stressed syllables, such as British English, as our consultants' exposure is almost exclusively to American English, which disallows [tju] in stressed syllables (McCarthy & Taub, 1992), as American English is the dialect taught in Brazilian schools, and finally, because even British English does not have [tju] for all of the words above. Finally, and most decisively, the spurious affrication of English [t] occurs only before [u], a restriction on the following vowel that could not be explained if it were only an attempt to approximate aspiration.

We propose that BPE speakers apply rules of their native BP to underlying forms they have set up. The application of Affrication (2) to the forms in (7) will thus yield spurious affrication.

The interest of this phenomenon is the fact that, as Peperkamp and Dupoux (2003) suggest, L2 and loanword phonology does involve an attempt to approximate the phonetic form of the donor language. However, in the case at hand, this

---

2.   Affrication with the voiced stop is rarer in our observations, but this may reflect our limitations as observers, since the fricative portion of [dʒ] is less salient than that of [tʃ].

approximation *is done through the phonology, and achieved by setting up an underlying form which contains the phonetic approximation.* Once this UR is set up, it is subject to the automatic rule of affrication just like any other underlying sequence of coronal plus high front vocoid.

3. Rhotic hypercorrection

In this section we report on the existence of BPE productions in which English *h*-initial words surface with an initial *r*. To understand why this puzzling and BPE-specific L2 phenomenon occurs, it is necessary to introduce first the distribution of rhotics within the native phonology.

3.1 Background on BP rhotics

We begin by reviewing the phonology of rhotics in BP. BP has three basic rhotics: [ɾ, ɦ, x], whose surface distribution is the following:

(8)   [ɦ] occurs syllable-initially when not postvocalic:[3]

    a.   *rabo* [ɦa.bu] 'tail'
    b.   *rei* [ɦej] 'king'
    c.   *roquenrou* [ɦɔ.kẽ.ɦow] 'rock and roll'
    d.   *honra* [õ.ɦa] 'honor'
    e.   *israel* [iz.ɦa.ɛw] 'Israel'
    f.   *dahruj* [da.ɦuʒ] (Proper name)

(9)   [x] occurs in the coda:[4]

    a.   *mar* [max] 'ocean'
    b.   *carne* [kax.ni] 'meat'
    c.   *circo* [six.ku] 'circus'

(10) [ɾ] occurs in complex onsets:

    a.   *prato* [pɾa.tu] 'plate'
    b.   *abre* [a.bɾi] 'open!'
    c.   *freio* [fɾe.ju] 'brake'

---

3. We assume that words like *honra* have an underlying nasal coda consonant that is deleted, following Mattoso Câmara Jr. (1970) and Wetzels (1997). This coda consonant conditions the allophony of the following postconsonantal [ɦ].

4. There is a wide range of sociophonetic and variationist work on the realization of coda rhotics in BP, which may be more velar, uvular, or glottal, depending on a variety of demographic and geographic factors; see, among others, Callou, Leite, and Moraes (2002). As our focus here is on word-initial rhotics, this variation lies outside the scope of the current problem.

(11)   [ɾ] and [x] contrast intervocalically:[5]

    a.   *carro* [ka.ɣu] 'car'
    b.   *caro* [ka.ɾu] 'dear'
    c.   *barra* [ba.ɣa] 'bar'
    d.   *barato* [ba.ɾa.tu] 'cheap'

The literature on whether the fricatives and the rhotic are simply distinct pho-nemes or are allophones in (near-)complementary distribution is vast. We adopt the view, following Mattoso Câmara Jr. (1953), Lopez (1979), Oliveira (1997), Mateus and d'Andrade (2000) and Abaurre and Sandalo (2003) that all of these surface allophones reflect a single underlying phoneme, which we posit is /r/.

    Many lines of evidence point towards this conclusion. The first comes from affixation and sandhi phenomena, which demonstrate that a coda [x] can become an [ɾ] when followed by a vowel.

(12)   a.   *por* [pox] 'through'
      b.   *por cima* [pox.sĩ.ma] 'through above'
      c.   *por aqui* [po.ɾa.ki] 'through here'

(13)   a.   *flor* [flox] 'flower'
      b.   *flores* [floɾes] 'flowers'

The converse occurs when an intervocalic [ɾ] becomes [ɦ] when truncation occurs:

(14)   a.   *direto* [dʒiɾɛtu] 'straight'
      b.   *reto* [hɛtu] 'straight, straight on'

We propose that the underlying rhotic is thus the tap [ɾ], which undergoes the following rules:

(15)   Onset debuccalization: r → ɦ in {#,C}._
(16)   Coda spirantization: r → x in _.{#,C}

Finally, we analyze intervocalic [x] as the result of a heterosyllabic geminate (see, e.g. Harris (1983, 2002)) for Spanish. Note that each half of the geminate will undergo one of the rules in (15) and (16).

3.2   BPE rhotic hypercorrection

We turn to the phenomenon of central interest in this section: the fact that BPE speakers occasionally produce tokens such as the following:

---

5.   Fricatives are subject to intervocalic voicing in BP, hence the allophone [x] undergoes a further change to [ɣ]. Regressive voicing in consonantal sequences also yields [ɣ] in words like *turma* [tuɣ.ma].

(17)   Sporadic hypercorrection of English *h*-initial words:

     a.   *home* [rom] ~ [hom]
     b.   *hug* [rʌg] ~ [hʌg]
     c.   *hunger* [rʌ̃gər] ~ [hʌ̃gər]
     d.   *hammock* [ræmək] ~ [hæmək]

The data on the left-hand side of the squiggle above are quite surprising at first blush. Why would a speaker who has heard the English word *home* pronounced as [hom] suddenly make the decision to pronounce it with an initial [r]? The descriptive answer is, this is hypercorrection. But what is the mechanism for hypercorrection, and how do the data in (17) fit within a broader theory of L2 and loanword adaptation?

We would like to make the proposal that *BPE speakers analyze incoming English words in terms of their native BP phonology, and set up underlying representations based on their native phonology*. In this respect our proposal resembles that of Ito et al. (2006), who propose that loanword adaptation involves the application of one's native inventory to the incoming data. In particular, we propose that the words in (17) have been set up with the URs in (18).

(18)   BPE speakers' URs for English words:

     a.   *home* /rom/
     b.   *hug* /rʌg/
     c.   *hunger* /rʌ̃gər/
     d.   *hammock* /ræmək/

Given the URs in (18), what accounts for the variable data in (17)? Our proposal is that speakers variably apply Onset Debuccalization (15) when speaking English. *The hypercorrected forms thus reflect the naked URs that these speakers have adopted.*

In fact, our proposal here resembles to some extent the assertion of Peperkamp and Dupoux (2003) that in L2 and loanword approximation, speakers attempt to match up the phonetics of the donor language as closely as possible. The difference is that the route to setting up this match is through the UR. BPE speakers who hear an English word-initial [h] indeed attempt to produce a surface [h] in their output grammars, but do so by setting up an underlying /r/ and letting the regular rules of BP underlying-to-surface mapping do the work of achieving the target output. Occasionally, however, these rules of UR-to-SR mapping fail, revealing the UR that BPE speakers have set up in their attempt to achieve the correct surface approximation of the English words.

A question arises at this point as to why the allophonic rule of Initial Debuccalization sometimes "fails" for the URs set up in (18). Why should the rule of Initial Debuccalization apply variably to these representations, when it applies categorically to native BP words?

The discussion thus far has only focused on English words that are surface *h*-initial, and we have not yet put forth a proposal as to the underlying representations of English words that are, in fact, surface *r*-initial, such as *radar*. Given the fact that BP has no possibility of a surface initial tap, what could the underlying representation be? Our proposal is that words such as *radar* are set up with an underlying tap, but are furthermore marked as exceptions to Initial Debuccalization. That is to say, on an item-by-item basis, these words must be diacritically annotated as exempt from the otherwise operative rule. Two representative grammatical models of exception-marking are Zonneveld (1978) in a generative phonology tradition and Pater (2006) in terms of lexically-specific faithfulness constraints in OT. For ease of exposition, we adopt the former here:

(19)   a.   *radar* /rejdar/ [– InitialDebucc]
       b.   *respect* /rispɛkt/ [– InitialDebucc]
       c.   *Redford* /rɛdfərd/ [– InitialDebucc]

There is, to a limited extent, already precedent for exception-marking to allophonic rules in the native phonology of BP. One example comes from truncation of words with intervocalic rhotics. The cocktail known as a *caipiroska*, made with crushed limes, has undergone a number of adaptations with different fruits, such as *morangoroska*, made with crushed strawberries. Much like the development of the English morpheme *-tini* (originally from *martini*, but now found in such coinages as *appletini* and *chocolatetini*), a truncated form *roska* has emerged as a catch-all term for cocktails such as *caipiroska, morangoroska, mangaroska*. Interestingly, this word is pronounced as [ɾoska], with an initial tap. Given that this word stands as an exception to Initial Debuccalization, it must have the UR in (20–c):

(20)   a.   *caipiroska* [kajpiɾɔska]
       b.   *roska* [ɾɔska], *[ɦɔska]
       c.   /rɔska/, [– InitialDebucc]

Exceptionally-faithful underlying forms resistant to allophonic rules can also be found with the otherwise extremely general and productive rule of Liquid Fronting in the plural, shown in (21):

(21)   a.   *jornal* [ʒoɣnaw, ʒoɣnajs] 'newspaper, SG. & PL.'
       b.   *radical* [ɦadʒikaw, ɦadʒikajs] 'radical, SG. & PL.'
       c.   *hotel* [otɛw, otɛjs] 'hotel, SG. & PL.'
       d.   *pastel* [pastɛw, pastɛjs] 'pastry, SG. & PL.'
       e.   *caracol* [kaɾakɔw, kaɾakɔjs] 'coil, SG. & PL.'
       f.   *sol* [sɔw, sɔjs] 'sun, SG. & PL.'
       g.   *anzol* [ẽzɔw, ẽzɔjs] 'hook, SG. & PL.'
       h.   *lençol* [lẽsɔw, lẽsɔjs] 'layer, SG. & PL.'

Liquid Fronting may be formulated as follows:

(22)   [+liquid] → [−consonantal, +high, −back] / +s#

Nonetheless, at least two nouns in BP (23) must be marked as [−LiquidFronting]:

(23)   a.   *gol* [gow, gows] 'goal, sg. & pl.'
     b.   *Skol* [skɔw, skɔws] 'Skol (brand of beer), sg. & pl.'

Thus, the existence of exceptional non-undergoing has ample precedent in the phonology of BP. Returning to the representation of *h*-initial vs. *r*-initial words in English, we may now contrast the two divergent underlying representations of *home* and *rome*, for example:

(24)   BPE speakers' URs for English words:
     a.   *home* /rom/
     b.   *Rome* /rom/ [−InitialDebucc]

As noted by Zonneveld (1978) and many others, the existence of exception-marking diacritics may be somewhat grammatically fragile. We propose that the existence of minimal pairs such as (24a,b) that differ only in the presence of a rule-non-undergoing diacritic are what lead to the occasional inhibition of Initial Debuccalization that gives rises to Rhotic Hypercorrection. Naturally, the occasional omission (or failure-to-access) such diacritics also may give rise to occasional BPE productions of words like *radar* with an initial [h].

Most crucially for the existence of Rhotic Hypercorrection, however, is the occurrence of a UR that is set up in accordance with the native phonotactics of BP in mind and subject to the normal rule of Initial Debuccalization, which will normally yield an output identical to the intended English word. That is, with an "incorrect" (or, more neutrally, *divergent*) UR, BPE speakers usually arrive at the correct target output for English *h*-initial words. Sporadic suppression of the allophonic rule producing this convergent output reveals the otherwise hidden divergent UR.

4.   Conclusion

Both spurious affrication and rhotic hypercorrection are puzzling phenomena from the perspective of either matching the surface phonetics or choosing the most identical underlying match for the donor language's correspondent. We have proposed a new model of L2 and loanword phonology here, one in which the speakers indeed attempt to match the surface forms of the donor language, e.g. English [tɕu] (or [tʉ]) and [hom], but must do so using the URs of their native

phonology, viz., /tʲu/ and /rom/. The crucial factor that yields spurious affrication and rhotic hypercorrection is the (sometimes abstract) form in which words are stored in long-term memory during an early stage of contact with the donor language. When the UR-to-SR allophonic rules of the native language are left to take over, the surface result may end up different from the input.

References

Abaurre, Bernadete, & Filomena Sandalo. 2003. Os róticos revisitados. In *Teoria Lingüística: fonologia e outros temas*, 144–180. João Pessoa: Editora Universitária da UFPB.

Boersma, Paul and Silke Hamann. This volume. Loanword Adaptation as First-Language Phonological Perception.

Cagliari, Luiz Carlos. 1997. *Fonologia do Português: Análise pela geometria de traços*. Campinas: Edição do Autor.

Calabrese, Andrea. 2005. *Markedness and Economy within a Derivational Model of Phonology*. Mouton de Gruyter.

Calabrese, Andrea. This volume. An Analysis-by-Synthesis Model of Loanword Phonology. Paper presented at Going Romance, Amsterdam.

Dinah Callou, Yonne Leite & João Moraes. 2002. Processo(s) de Enfraquecimento Consonantal no Português do Brasil. In *Gramática do Português Falado VIII*, 537–555. Campinas: Editora da UNICAMP.

Hall, Tracy, & Silke Hamann. 2006. Towards a typology of stop assibilation. *Linguistics* 44.6:1195–1236.

Harris, James W. 1983. *Syllable Structure and Stress in Spanish*. MIT Press.

Harris, James W. 2002. Flaps, Trills, and Syllable Structure in Spanish. In *Phonological Answers (and their Corresponding Questions), MITWPL 42*, 81–108. Cambridge, MA: MITWPL.

Ito, Chiyuki, Yoonjung Kang, & Michael Kenstowicz. 2006. The Adaptation of Japanese Loanwords into Korean. In *Studies in Loanword Phonology, MITWPL 52*, 65–104. Cambridge, MA: MITWPL.

Kim, Hyunsoon. 2001. A phonetically based account of phonological stop assibiliation. *Phonology* 18:81–108.

LaCharité, Darlene, & Carole Paradis. 2005. Category Preservation versus Phonetic Approximation in Loanword Adaptation. *Linguistic Inquiry* 36.2:223–258.

Lopez, Barbara. 1979. The Sound Pattern of Brazilian Portuguese (Cariocan Dialect). Doctoral Dissertation, UCLA.

Mateus, Helena, & Ernesto d'Andrade. 2000. *The Phonology of Portuguese*. Oxford University Press.

Mattoso Câmara Jr., Joaquim. 1953. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Organizações Simões.

Mattoso Câmara Jr., Joaquim. 1970. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes.

McCarthy, John, & Alison Taub. 1992. Review of Paradis & Prunet (eds.), The Special Status of Coronals. *Phonology* 9:363–370.

Oliveira, Valéria. 1997. Um reestudo da vibrante: análise variacionista e fonológica. Doctoral Dissertation, Pontifícia Universidade Católica do Rio Grande do Sul.

Pater, Joe. 2006. The Locus of Exceptionality: Morpheme-Specific Phonology as Constraint Indexation. ROA-866.

Peperkamp, Sharon, & Emmanuel Dupoux. 2003. Reinterpreting Loanword Adaptations: The Role of Perception. In *Proceedings of the 15th International Conference of Phonetic Sciences*, 367–370.

Wetzels, Leo. 1997. The Lexical Representation of Nasality in Brazilian Portuguese. *Probus* 9.2:203–232.

Zonneveld, Wim. 1978. *A Formal Theory of Exceptions in Phonology*. Peter de Ridder Press, Lisse.

Early bilingualism as a source of morphonological rules for the adaptation of loanwords

Spanish loanwords in Basque*

Miren Lourdes Oñederra
University of the Basque Country

The present socio-cultural situation in the Basque speaking area of Spain offers a privileged field for the study of Spanish loanwords in Basque, due to the more expanded use of Basque, together with a better knowledge of Spanish among Basque speakers. Within the theoretical framework of Natural Phonology, this paper explores some phonological and *lato sensu* morphological mechanisms that take part in the integration of Spanish loanwords into Basque. First it deals with the mutual influence between Spanish and Basque when both are first languages for the speaker. Early bilingualism only causes the loss of Basque processes that are suppressed in Spanish, but those processes need not be completely lost. There is clear evidence that continued collective bilingualism and need of translation motivate the transformation of denaturalised phonological substitutions into morphological devices for the adaptation of loanwords.

1.   Phonological influence in early bilingualism

This paper deals with the pronunciation of Spanish-Basque (or vice versa) bilingual speakers of the Autonomous Community of the Basque Country (ACBC). This area offers the most appropriate setting for our study, given its present socio-linguistic situation. It should be noted that Basque and Spanish are both official

languages in the ACBC, and therefore a relatively high degree of collective and individual bilingualism can be found among its inhabitants. Some of them learn Basque as a second language; others acquire Basque and Spanish during childhood. The relatively extended knowledge of Standard Basque is also an important factor in the configuration of the present linguistic situation.

There is a further reason for me to have chosen this area: it is the one I know best, and I take advantage of this opportunity to render my little homage to Kruszewski by quoting his words here. As will be shown later, the basic distinction established in that publication by Kruszewski between different types of alternations is fundamental to the theoretical views underlying this paper.

> The German reader would certainly have found this publication much more convincing had I selected German examples. To do so, however, would have required complete competence in colloquial German, which I cannot claim. I was therefore obliged to resort to examples from Polish, my native language, and from Russian, in which I am fluent. (Kruszewski 1978:64)

I furthermore think that direct experience –both sociological and linguistic– is of great import at the present stage of analysis. The study of how the relationship between the two languages develops in our community will be carried along the lines of the theory of Natural Phonology (NP) as proposed by David Stampe (1969, 1979; Donegan & Stampe 1979). Indeed, the subject came up in the process of preparing a book on NP (Oñederra in prep.) applied to Basque, specifically from projection of a concept fundamental to NP, the concept of *phonological process* onto the bilingual context, particularly one type of bilingualism which will –for lack of a better name– be called *close bilingualism*. 'Close' is here meant to include the notion of *early* bilingualism, as far as individual development of the speaker is concerned, as well as language contact *continued* over centuries in the collective history of the community. That collective historical dimension will be shown to be essential for the hypothesis presented here.

A phonological process as conceptualised in NP is "a mental operation that changes a given segment or sequence that presents an articulatory or perceptual difficulty into another segment or sequence that lacks that difficulty" (Hurch 1988:350). I will try to show how this is a useful tool that may enable us to predict when phonological interference of one language over the other should take place in bilingual speech, and what the (phonetic) shape of such an interference would be (see Section 2 below).

The NP concept of phonological process will allow us to diagnose (a) the degree of bilingual competence of individual speakers, and (b) the general state or productivity of Basque sound substitutions. In other words, the analysis of the speaker's active phonological processes will measure how robust the phonological

system of each language is, given what can be expected when two languages have been acquired in early childhood. When phonological competence is not even, precedence of one language over the other should be detected for each pronunciation phenomenon observed (see different situations in Section 2).

In Section 3 the concept of *morphonological rule* developed by NP will be the complementary theoretical resource to account for a phenomenon which is particularly productive in Basque nowadays: the use or 'recycling', as translation rules in the adaptation of loanwords from Spanish, of processes which have lost their phonological status (see Section 3 below). Although that loss of phonological status (i.e. phonetically motivated productivity) may or may not be due to the influence of Spanish, this paper will focus on those cases where early bilingualism seems to play a fundamental role.

From the NP perspective, a rule is phonetically conventional and "distinct in its nature, evolution, psychological status and causality" (Donegan & Stampe 1979:127) from a process. Rules may take on morphological motivation, though that is not necessary. Among the phonetically conventional sound alternations Kruszewski (1978:70,73) distinguishes alternations without a morphological function whatsoever (The Second Category) from those that may be linked to such a function (The Third Category). But, as far as the explanatory realm of phonology reaches, the fundamental limit lies between phonology and grammar, as solidly established by Sapir or Kruszewski (see Donegan & Stampe 1979:127), and now developed by NP.

This paper argues that the ACBC bilingual setting provides a particularly appropriate atmosphere for a given type of morphological rules to develop, which will ultimately be used to change the phonological configuration of Spanish words borrowed by Basque.

Sound substitutions that have lost their phonetic motivation, but nevertheless remain in the language, may eventually become such morphonological rules. This transformation can be analysed as a sign of linguistic health in the bilingual setting which, in turn, would demonstrate the functionality –and hence productivity in synchronic terms– of a certain type of morphonological rule that would develop in some situations of 'conscious' bilingualism. In that sense, this work may contribute to the development of some theoretical elements of NP by their application to the bilingual scenario.

The term 'morphonological' (i.e. morpho-phonological) may be senseless, at least from a purely functional point of view. As will be claimed in Section 3, sooner or later these rules become functionally equivalent to morphological suffix-correspondence rules, only that, as a consequence of the fact that they result from originally *phonological* differences between the two languages, these rules tend to produce phonologically more similar pairs of words than other correspondence

patterns based on lexical differences (like, e.g. adjective-forming Spanish *-ble* > Basque *-garri*, or verb-forming Spanish *-r* > Basque *-tu*). In fact they often are a means of translating what may be analysed as a suffix in Spanish by what is considered its Basque counterpart, phonologically similar but different enough (due to the correspondence pattern established by the sound substitution). The alleged suffix need not be so grammatically (see in Section 3 examples of Spanish -ón as in *botón* "button", *futón* "futon", etc.), but are, as Picard and Nicol (1982:165) would say, "psychologiquement réel et morphologiquement productif ".

The main reason to keep the term morphonogical is the wish to underline the frequent phonological origin of these rules. It should also be added that, following Stampe's NP, there is a clear-cut categorical distinction between phonological and non-phonological phenomena: once a substitution has lost its phonetic motivation, it is not part of phonology proper. So calling it morphonological or simply morphological is, at most, secondary.[1]

Besides, 'morphonological' is particularly adequate referred to the phenomenon of sibilant affrication in Basque, which will be the main illustration of the ideas proposed in this paper (e.g. Spanish *consigna* > Basque *kontsigna*, see Section 3). Indeed it will be proposed that the change from process into rule of post-sonorant affrication is happening at present. Therefore phonetic motivation is still quite transparent. It might even be the case that it still keeps its phonological status for some speakers, while it has already become morphological for others. On the other hand, it is not possible to analyse affrication as a suffix correspondence rule, since the fricative-affricate substitution is stem-internal and cannot be segmented as a suffix.

2.   Early bilingualism: Denaturalisation of processes?

Close bilingualism refers to two first languages, i.e. when two languages are acquired more or less simultaneously.[2] In order to limit the period of acquisition somehow, however approximately, we can say that the speaker must have enough exposure to the two languages during her/his phonological formation, that is, the period during which the language specific options she or he is acquiring have not yet become a perceptual or productive constraint. Since that transformation from option into

---

1.   See the semininal work by Dressler (1985) for a different (gradual transition from phonology to morphology) proposal within the NP paradigm.

2.   We could also call it native knowledge of both.

constraint seems to be over by adolescence, both languages should have been acquired before then.

> Gradually we constrain those processes which are not also applicable in the mature language (…). From adolescence usually there is little further change, and the residual processes have become the limits of our phonological universe, governing our pronunciation and perception even of foreign, invented, and spoonerized words, imposing a 'substratum' accent on languages we subsequently learn, and labeling us as to national, regional, and social origins. (Donegan & Stampe 1979:126–127)

At present, practically every individual who speaks Basque as a first language has also acquired Spanish before adolescence in the ACBC. The same could be said about Basque and French in the Basque speaking area of France. Due to the different phonological characteristics of French, that area offers a very interesting point of reference and contrast indeed, which should be taken up by future research.[3]

In order to proceed in the analysis of the double phonological acquisition in the ACBC, let us now return to the concept of *phonological process*, and consider the theory of NP, where processes are

> (…) mental substitutions which systematically but subconsciously adapt our phonological intentions to our phonetic capacities, and which conversely enable us to perceive in others' speech the intentions underlying these superficial phonetic adaptations. (Donegan & Stampe 1979:126)

A phonological process is, therefore, a mental substitution that responds to a phonetic, i.e. physical, difficulty related to the articulation or perception of segments and sequences. Those difficulties are *per se* universal and so are the natural processes that eliminate them. But not all languages retain the same processes in their phonological systems. In short, for those who may not be familiar with the theory, it is a language-specific option:

a. to avoid a given phonetic difficulty (e.g. context-free vowel nasalization) by *allowing* a certain process to apply (i.e. vowel denasalization), or
b. to *overcome* the process by learning how to pronounce and perceive the difficult phonetic configuration. That segmental or sequential configuration will then become part of the phonology of that given language where the phonological process (that would have relieved the speaker from having to cope with the difficulty) will no longer be present. In other words, at a given moment during

---

3.   The significant socio-political differences would also add to the interest of such an investigation.

the speaker's language acquisition, that process will disappear from her/his actual and potential competence.

If the eliminated process is a context-free process (that would have avoided a segmental difficulty), the language adds one phoneme or phoneme class to its inner phonemic inventory. If the eliminated process is a context-sensitive one (that would have avoided a sequential difficulty), the language will have one more possible sound sequence (that speakers will be able to distinguish and intentionally pronounce). If the process had prevailed, that sequence would be excluded.

Consequently, therefore, for each phoneme or phoneme class that is acquired a context-free paradigmatic process must be eliminated. So, French must have overcome the universal process V → [–nasal] of vowel denasalization (motivated by the phonetic optimality –better articulatory and perceptual quality– of oral vowels), in order to have both oral and nasal vowels in its phonemic inventory; Basque must have overcome the universal process [+strident] → [+ cont],[4] in order to have both phonemic fricative and affricate sibilants.

In the same way, each new acquired sequence brings about the elimination of a context-sensitive syntagmatic process. Languages with voiceless intervocalic obstruents are a clear example: the universal phonetically motivated process of intervocalic obstruent voicing must have been overcome by their speakers in order to be able to produce sequences of *vowel-voiced obstruent-vowel* vs. *vowel-voiceless obstruent-vowel*. Put simply, languages that allow the process to apply will only have *vowel-voiced obstruent-vowel* sequences (e.g. S. Chinook or Sanskrit, see Donegan 1995:64–65).

Processes are not borrowed as such.[5] How could they be since they are universal? What one language may borrow from another is the elimination of a given process. That is a language specific option, which may be the source of differences between two languages and the cause of interference in close contact situations. It is not hard to believe that, if two languages co-exist during acquisition, any phonetic, i.e.

---

4.   The exact name of the feature chosen is not so important here. It may be more useful descriptively to talk about continuant sibilants, instead of the Jakobsonian labels used. The ultimate goal in the choice of labels would be to be as close as possible to the most plausible phonetic explanations for the substitution at hand. Experimental phonologists could be called on for help here.

5.   There might be activated latent processes that may look like borrowed processes. See Churma 1984:226, Hurch 1988b, on latent processes. See also Calabrese (this volume, "Acoustic inputs and phonological discrimination") about crucial age of exposure to the most discriminating language.

physical, difficulty overcome by a speaker in the phonological acquisition of one of her/his two languages will become an ability of that speaker, a perceptual and articulatory resource of her/his linguistic competence, available to her/him when facing the task of perception and production of the other language. The process of acquisition can be seen as a series of changes in the sound pattern of a speaker, which recalls Donegan (1993:98): "sound changes are changes in the speaker's phonetic abilities".

In other words, when a speaker has enough phonological command of two languages LA (language A) and LB (language B) due to early enough acquisition of both, different specific choices within LA and LB when faced by the same phonetic difficulty may result in conflict. But, given that a process is the realization of a phonetic limitation, the reflection of a physical difficulty, a certain pattern can be predicted for the resolution of that conflict. For instance, if LA allows a natural process X → Y to apply in order to solve the phonetic problem, but LB overcomes the difficulty by eliminating the facilitating process X → Y from its phonological system (where X will be integrated), the speaker who has acquired phonologies A and B will be able to overcome difficulty X, and will not *need* to apply the process X → Y either in LB (where X exists normally) or in LA. LA may keep the process as an optional more or less productive substitution. A good example of this is the affrication of sibilants following sonorant consonants in Basque (e.g. *pentsatu* "to think"), non-existent in Spanish (cf. *pensar* "to think") and not anymore a necessity in the pronunciation-perception of Basque-Spanish bilingual speakers (i.e. pattern (a) below).

At this point it may be worth giving some thought to the fact that if acquisition is bilingual, phonological transfer from one language to the other will not depend on sociological language dominance, but on the actual process-share of each of the phonological systems in contact. Some specific cases from the Basque-Spanish contact will illustrate three possible patterns of process distribution between the two languages:

a. when Basque keeps a process that Spanish does not allow,
b. when Spanish keeps a process that Basque has overcome, and
c. when both languages keep a process.

Situation (a) Basque keeps a process that Spanish has overcome. For example, affrication after sonorant consonant applies in Basque, but not in Spanish (cf. Basque *pentsatu*, Spanish *pensar* "to think" from Latin *pensare*). It is clear that the articulatory capacity to pronounce sibilant fricatives after a sonorant consonant, which an early bilingual speaker must have mastered in order to cope with the phonology

of Spanish, enables him or her to also pronounce them in Basque. Therefore the process of *sequenced* nasal-sibilant transition of Basque phonology becomes at most an optional substitution. It follows automatically therefore, that any Basque process that does not have an equivalent counterpart in Spanish will not be an obligatory phonological process in the linguistic competence of an early Basque-Spanish bilingual speaker. [6]

A paradigmatic context-free example of the same situation (Basque applies a process, but Spanish does not) can be observed in the process whereby all coronal fricatives are turned into sibilants in Basque, while this process is absent in Castilian Spanish.[7] This leaves Castilian Spanish with a basic contrast between /θ/ and /s/ (cf. *casa* [kasa] "house" vs. *caza* [ka θa] "hunt"), which is absent from the phonemic inventory of Basque. However, if Spanish is acquired early enough, the process will be eliminated, and bilingual speakers will be perfectly able to perceive and intentionally produce (see modern loanwords like *pro*[θ]*esu* "process", *so*[θ]*iologia* "sociology", [θ]*entro* "center" of bilingual speakers who are not constrained by the Basque process any more).

Situation (b) Spanish keeps a process that Basque has overcome. This situation can be illustrated by the paradigmatic context-free choice that links palatality of obstruents with affrication in Spanish, but not in (most dialects of) Basque. As a consequence of this, the only palatal obstruent in the phonemic inventory of Spanish is /tʃ/, whereas the Basque inventory adds /c/ to /tʃ/.

Basque phonology of early bilinguals will not be affected under these circumstances. The phonology of the language is immune to interference in situation (b), no matter how strong the sociological influence of Spanish might be. Once a process has been overcome, nothing should be able to reactivate it as such a process. What is crucial here is the degree of phonological competence (i.e. phonetic command) that is usually guaranteed by early acquisition. At any rate, if the Basque pronunciation of bilingual speakers shows traces of this interference, this will be a sign of asymmetry, showing that Spanish has taken precedence over Basque during acquisition. The truly early bilingual will have a larger inventory of phonemes (ergo perceptual discrimination capacity) and/or more types of sound sequences will be possible for her/him.

Examples of this type are easy to find in the Basque-Spanish setting among paradigmatic context-free processes. As well as the palatal stop that is added to the affricate in Basque, we find that the Basque phonemic inventory, on top

---

6. Its first stage could easily consist in becoming a regular process systematically applying to native forms, but only optionally to loanwords and neologisms. It could then become a morphonological rule (see below, Section 3).

7. Castilian is the Spanish variety spoken by Basque-Spanish bilinguals in the ACBC.

of distinguishing three fricative sibilants (apical alveolar, laminal alveolar and pre-palatal), includes affricate counterparts of all of them. Being perfectly competent in Spanish does not do any harm in this case. Bilingual speakers have both affricate and fricative (laminal and apical) sibilants in their phonemic inventory, or both affricate palatals and palatal stops. Neither the latter nor the affricate sibilants will be used in their pronunciation of Spanish native forms, but they may be helpful when ordering a *pizza*, with which monolingual Spanish speakers have some difficulty ([pitʃa], [pisa], [piθa] are common among monolingual speakers).

As said before, in order for Basque to be immune to Spanish interference under the circumstances characterized as situation (b), both languages with their whole phonological systems must be, so to speak, first languages (L1), in actual fact and not only apparently, ideologically or intended. In other words, if a bilingual speaker acquires Spanish as her/his L1 (i.e. acquires phonic command of Spanish) and only learns Basque some time later (when phonetic options have become phonological limitations), her/his pronunciation of affricates or the palatal stop in Basque may be affected: she/he will tend to make sibilants always fricative, or to reduce affricates to palatal [tʃ]), will substitute /tʃ/ for Basque /c/.[8]

Situation (c) The third possible pattern is that in which neither Basque nor Spanish overcome a phonetic difficulty, i.e. both Basque and Spanish allow one given process to apply.[9] All the processes shared by both languages are to be included here, like the context-free denasalization (V → [–nasal]) that explains why Basque and Spanish only have oral vowels in their phonemic inventories. Among context-sensitive processes, a good example is intervocalic assibilation of voiced stops, which is productive in both languages.

When the two languages share a process, no change can be induced by any of the two languages onto the other one. So, we can say that phonological interference from a given language B (LB) on language A (LA) in early bilingualism will only happen when LB has eliminated a process that is active in the phonology of LA (situation (b) above). However, given that eliminating a process implies having overcome a given phonetic difficulty, giving up the process not shared by LB will

---

8.   This is a particularly interesting subject for comparison with the production of French-Basque bilinguals, as French lacks /tʃ/. Of course, late acquisition and early acquisition of incomplete inventories from parents or teachers may yield similar results.

9.   The latter formulation is actually more precise. Things would probably be different if Basque and Spanish would admit a certain difficulty by making use of two different processes. The relatively common choices of Basque and Spanish allow me to avoid –for the time being– this issue which, however interesting, falls outside the reach of this paper. At first sight, it would appear that no problem (interference) should arise: cf. English *ss* → *ses (fishes)* vs. Basque *ss* → *ts*; Spanish d# → (∂) → ø, d# → θ vs. Basque d# → t. All of these, however, should fall under either situation (a) or (b), and they must be analysed separately.

bring about the gain of new phonemic units or sequences in LA. Speakers will have mastered the phonetic difficulty that the process avoided, being consequently more capable in terms of phonological productivity and perception.[10]

In general terms, the prediction would be that the more *natural* LA is (i.e. the more natural universal phonetic processes LA keeps active in its phonology) and the less natural LB is, more change early bilingualism should bring to LA. Processes that LA could have kept in isolation (or when LB is only learned at best as a second language) will become optional for early bilingual speakers, or they will disappear completely from their phonology. In other words, a more 'elaborated' phonology (a phonological system that distinguishes more units or sequences, because of having overcome more natural processes) is more 'harmful' in terms of phonological influence on a less elaborated or more natural phonological system, because it will raise more instances of situation (a).[11]

3.   Metamorphosis: From pronunciation to translation

Once a given substitution X → Y has been liberated from its phonetic conditioning in LA, due to mastery of the corresponding phonetic difficulty during acquisition of an LB that does not apply process X → Y, what we have is a change from input X to output Y available to other possible linguistic functions. Put it in a different way, the substitution is no longer something that makes X (better) pronounceable or perceptible, it is not phonetically necessary any more in LA. The ability reached in the acquisition of LB makes the substitution phonetically superfluous. However, as we often see, some of these substitutions may well change qualitatively, and become morphonological rules, which may then acquire a grammatical or lexical function. I want to argue here that in a bilingual society, one of the possible functions of such rules is that of obtaining Basque forms from originally Spanish words (e.g. see below Basque ns → nts, as in Spanish *consigna* > Basque *kontsigna*).

---

10.   Careful distinction must be made in the description of the two languages between non-existent (but possible) sounds or sequences and the impossible ones, eliminated by active processes as discussed by Pensado (1985/1999).

11.   The Research Project (UPV05/81) now being undertaken by musicologists and linguists from the University of the Basque Country pursues the goal of a deeper understanding of Spanish and Basque prosodic patterns that may well be the fundamental basis of these two languages sharing so much of their phonologies. The comparison with French-Basque bilingualism is of great importance at this point, since French overcomes more natural processes than both Spanish and Basque.

A plausible requisite for that to be true is that the substitution should be phonemic (i.e. perceptible and memorizable by the speaker). Another important factor contributing to the metamorphosis at hand seems to be the productivity of the substitution in LA. That, together with a long history of permanent language contact, will increase the probability of parallel but crucially different cognates in the two languages. Then speakers may feel the substitution as 'proper pronunciation' in LA, even after phonetic motivation and justification is lost due to early bilingualism.

The following necessary characteristics are now present in the Basque-Spanish speaking community in the ACBC:

a.   A long history of continuous language contact, which helps to develop and consolidate patterns of cognates for words stemming from common etymos.
b.   A more expanded and complete knowledge of Spanish among Basque speakers, who are now practically all bilinguals. A gradual increase has occurred in the class of people who would in previous generations have had a reduced competence in Spanish, but who have a full command of it now. Other speakers were Basque monolinguals and are now Spanish-Basque bilingual speakers.[12] Early bilingualism has also increased, which strengthens the chance for Basque processes not shared by Spanish to lose their phonological productivity.
c.   The recent increase of Basque among new learners, as well as (very importantly) its expansion to new linguistic areas due to the officialization of the language, and the subsequent need for the urgent translation of Spanish words. This enhances the chances of processes becoming (translation) rules.

All these sociological factors create the motivation for the above mentioned metamorphosis: the phonological processes that cease to be so and are then free to take on other functions, become morphonological rules of loanword adaptation for bilingual speakers and a possible model or reference for the not so early ones.[13]

Continuous bilingualism is (and has long been) motivation for the productivity of rules derived from denaturalised processes (i.e. substitutions that are no longer

---

12.   Causes for this expansion of Spanish have been externally imposed or voluntarily adopted. It would of course take longer than is either possible or necessary here to say all there is to say on this from a sociolinguistic point of view.

13.   Pronunciations of not so early bilinguals like *deskan*[tʃ]*oa* (for *deskan*[ts]*oa*) from Spanish *descanso* (vs. vernacular *atsedena*) show that the rule of affrication may be used, even by speakers who lack the appropriate affricate (since, as a result of Spanish precedence in the configuration of their phonemic inventory, they never eliminated thoroughly the process reducing affricates to /tʃ/).

phonetically motivated) as well as of rules that have a different source. For example, vowel prothesis before word initial trills stems from an old phonological process attested to in Basque, common to all dialectal varieties. Vowel prothesis before a word initial trill is now a rule productive only in the adaptation of some loanwords, like Spanish *radical* > Basque *erradikal* "radical, extreme", Spanish *reseña* > Basque *erreseina* "review". Apart from this, trill initial forms can easily be found in Basque at present (e.g. *radar*, *radikal*, etc. among less purist speakers; cf. also the common form of proper names like *Ramon*, *Rosa* vs. old or literary *Erramun*, *Errosa*).[14]

Other correspondences also exist due to the different phonological choices of each language. Some of them are productive as translation rules, among others: Spanish (but not Basque) final [e] elision in Latin forms has yielded pairs like Spanish *amor vs.* Basque *amore* from Latin *amore(m)* "love" (Michelena 1995:146).

Lack of final [e] elision in Basque together with Basque (but not Spanish) inter-vocalic nasal deletion, followed by vocalic quality change and desyllabification, pro-duced pairs like Spanish *león* vs. Basque *leoi* [leoj] from Latin *leone(m)*.[15] Although both nasal deletion in Basque and final [e] elision in Spanish ceased to be part of phonology a long time ago, their diachronic results productively apply in the adap-tation of new loanwords, like Spanish *botón* "button", *camión* "truck", *futón* "futon", *evaluación* "evaluation" > Basque *botoi, kamioi, futoi, ebaluazioi*;[16] on the other hand Spanish *contestador* "answering machine", *radiador* "radiator" > Basque *kontesta-dore*, *radiadore* show the productivity of a synchronic rule of final [e] epenthesis which follows the pattern of that final [e] that Basque kept but Spanish deleted.

Intervocalic voicing, which took place at a certain stage in the history of Spanish (but not Basque), also belongs here. Voicing, followed by the already mentioned

---

14.   Previous analyses of vowel prothesis have been carried out under different theoretical assumptions. But, whatever the differences are among authors, they never account for the qualitative distinction between diachronically attested and living phonological patterns of the language. See among those who acknowledge that prothesis is not phonologically productive at present (still listing it as part of the phonological characteristics of Basque), Hass (1992:36). Similarly Trask (1997:146) points out the acquisition of word-initial [r] in loans, but no further consequence is derived.

15.   Hualde (2000:349) offers a detailed summary of the phonological evolution of those words in Basque and Romance.

16.   The latter is normatively wrong, as older Latin *-tione* endings correspond to Basque *-zio*. It has been included here, because this overgeneralization can be seen as a clear proof of the productivity of these rules in their new non-phonological domains. And also because it overrides Hualde's interpretation, according to which the Spanish *-ón* > Basque *-oi* change is bled by "the more specific rule, which reflects a correspondence between suffixes" (Hualde 2000:349).

final [e] elision, underlies old cognates like Spanish *universidad* vs. Basque *unibertsitate* from Latin *universitate(m)* or Spanish *virtud* vs. Basque *birtute* from Latin *virtute(m)*. These pairs must have set the pattern for the nowadays productive devoicing of equivalent suffixes being translated into Basque, like Spanish *idoneidad* (a certain type of qualification for jobs at the university), *titularidad* "tenure", *oportunidad* "opportunity" becoming Basque *idoneitate*, *titularitate*, *oportunitate* (occasional for vernacular *aukera*).[17]

On the other hand and back again to patterns emerging from processes that Basque has applied but Spanish has not, we find the affrication of sibilants following sonorant consonants. This affrication is particularly interesting, since it is a sound substitution which is changing from process into rule in the present day. Cognates result from an evolution that had already begun when Latin forms were adopted along different phonological paths in Spanish and Basque: cf. Spanish *oso* from Latin *ursu(m)* "bear", Basque *(h)artz* "bear" (from Latin *ursu(m)*?, cf. at any rate Aquitainian *Harsus*). Some correspondences must stem from the time when the present Castilian Spanish interdental was still a sibilant (i.e. before the 16th or 17th centuries, Cano 2004:843). It is clear that, for example, Basque *dantza* [dantsa] "dance" was not 'phonologically' derived from Spanish *danza* pronounced [danθa]. But nowadays [ts] substitutes for [θ] as a result of the systematic translation rule that productively changes Spanish *pinza* [pinθa] "tweezers", *trance* [tranθe], *sentencia* [sentenθja] "law sentence", etc. into Basque *pintza*, *trantze*, *sententzia*, etc., phonetic opacity being an exclusive characteristic of rules (vis-à-vis phonetically motivated processes). As a matter of fact, one source of affrication may already be present in some stages of the Romance evolution.[18]

In order to develop an analysis of the present situation of Basque affrication of sibilants following sonorant consonants, let us focus on examples of its synchronic application to Spanish loanwords like *consigna* "slogan", *corresponsal* "correspondent

---

17.   Whether these word-endings are analysed by speakers as suffixes or we are dealing with word-adaptation patterns would be an interesting subject, but it is beyond the scope of this paper. In any case, it may be worth noting that these endings are never attached to non-borrowed stems. See Oñederra 2002, for a more complete, though by no means exhaustive, list of this type of rules.

18.   As seen before, rules are often old processes which have become obsolete, and are not to be synchronically explained by phonetic characteristics of the language (e.g. Spanish intervocalic voicing attested by Latin *-tate* becoming *-dad*); other rules accumulate the effect of several processes which have diachronically fed the present result (e.g. nasal deletion, vowel change and desyllabification in Latin *leone(m)* > Basque *leoi* [leoj]. Following Kruszewski (1978:70), "The causes or conditions of such an alternation can only be discovered by investigating the history of the language", and they can even be "completely unknown" (Kruszewski 1978:74).

(reporter)" and *insumisión* "movement against military service", which have become Basque *kontsigna*, *korrespontsala*, *intsumisio* (also *intsumisioi*, see Fn. 16). We should bear in mind the basic notions of the theory of NP recapitulated here:

a. Changes in phonology are changes in the phonetic capacities of speakers.
b. There is a clear-cut boundary between phonology and morphology, drawn by the interplay between the phonetic motivation of the phonological dimension and its absence on the morphological side of the boundary.
c. Phonological substitutions can cease to be so (ontologically, so to speak) and become part of morphology. That is known at least since Kruszewski taught at the end of the XIX century (see Kruszewski 1978) and Wurzel (1980) developed the idea within the NP framework. In this paper we go on to imply that those substitutions might become part of morphology as lexicon formation devices.

From the point of view of NP, the affrication of sibilants after sonorant consonants in itself may perfectly be seen as a phonological process, a constraint, a need that responds to the phonetic complexities of nasal-sibilant or liquid-sibilant transitions.[19] As a process, it allows present day Basque speakers to avoid the complex transition in *kontsigna* "slogan", *boltsa* "bag" or *kurtso* "*course*" (from Spanish *consigna*, *bolsa*, *curso*), or in alternating Basque forms like the auxiliary verbal form at the end of *ekarri[s]uen*) "he/she brought (it)" vs. *esan[ts]uen* "he/she said (it)", in the same way that earlier generations did when they first pronounced *(h)artz* "bear", *faltsua* "false" or *dantza* "dance". But if speakers are confronted with the phonology of Spanish, which does not allow the facilitating process to apply, in early childhood (that is, before their phonetic abilities become their phonological limitations, when abilities can still grow and expand with no effort) the relatively more difficult sonorant-sibilant consonant sequence will be learned. Substituting an affricate will no longer be a phonological need for that speaker. This is why sibilant affrication no longer forms part of the phonology of most Basque speakers nowadays. There is no phonological trend in the substitution. It happens in words in which it has been lexicalised and so learned when learning the inner representation of the lexicon. Only speakers who still have alternations in morpheme

---

19.   Busà (2007) offers an excellent review of the phonetic and phonological work done so far on nasal-sibilant sequences, together with an interesting comparative work on Italian affrication (vs. English), and promising paths for future research. I wish to thank Maria Josep Solé for letting me know about this paper, which was still in press, when we met at the PAPI 2007 conference in Braga. Jauregi & Oñederra (2008) explores phonetic and phonological characteristics of the liquid-sibilant sequences in Basque.

boundaries may be said to keep the process (although only optionally in most cases). The systematic application of the substitution to loanwords would be the consequence of its new status as a morphonological rule of loanword adaptation (through the morphological function of forming *Basque* words).

As pointed out before, certain features are constant in this type of transformation from phonological processes into morphonological rules for the adaptation of loanwords:

a. Initially, different choices of Basque and Spanish phonologies for a given phonetic difficulty which can be observed in the morpheme internal consonant sequences of native forms, i.e. sonorant-fricative sequences are perfectly regular in Spanish (*pensar* "to think"*, cansado* "tired", *cursi* "ridiculous", *pulso* "pulse"), but impossible in Basque (***anza, **elze, **hersi* vs. *antza* "similarity", *eltze* "pot", *hertsi* "close").
b. Next, bilingual speakers become conscious of the fact that there are correspondences between similar but slightly different forms in Spanish and Basque (cognates like *pensar/pentsatu*).
c. Finally, speakers use the relative difference as a means to translate Spanish words into Basque, taking a substitution which is no longer part of their phonetic limitations as the basis for the nativization of Spanish forms. [20]

As far as the phonological analysis is concerned, it is important to note that the process ceases to be an obligatory process of Basque phonology, if it does not disappear altogether. That is, *close* bilingualism does not perhaps mean the immediate loss of a process in Basque, because of Spanish acquisition, but the process will become optional, and therefore weaker. Once that occurs, bilingualism will intensify and speed up the transformation of the substitution into a morphonological rule, which will then be functionally motivated as a translation rule.

4. A first provisional conclusion

Although loss of phonetic motivation and, therefore, phonological status of a given substitution may be caused by bilingual acquisition of individual speakers,

---

20. Hualde (in Hualde & Ortiz de Urbina 2003:62), though clearly stating that "In this adaptation one can see a conscious attempt to preserve aspects of the traditional phonology of the Basque language", still considers affrication part of the synchronic Basque phonology. That of course is what could be expected from his theoretical standpoint, in which phonetic motivation is not a structural property of phonology.

productive use of that substitution as a translating device will require some collective support.

On the other hand, our incipient study on Basque shows that enough early bilinguals who master the phonological system of the language are necessary so that Basque phonological alternations survive, even if they will no longer be sustained by phonological processes but used as a way of adapting Spanish loanwords. At least in the period when the transformation of a process into a rule takes place in the language, only if enough speakers keep the substitution as part of their first language acquisition, will a productive pattern of correspondence between the two languages be structured. From then on the substitution can productively survive as any other rule of loanword adaptation (examples of those rules were given in Section 3).

It is clear that bilingualism can be the reason for the weakening and eventual loss of phonological processes in one language, whenever those processes are absent in the other language simultaneously acquired. It is clear that, from a strictly phonological point of view:[21]

> The most interesting interference phenomena attested by loanwords come to light when the speakers of L_T, who borrow from L_S are nearly monolingual, or when these mediators are imitated by monolingual speakers of L_T with no attempt to adjust their speech habits to the phonology of L_S. We may expect less evidence of extreme interference proportional to the greater degree of bilingualism of the borrowers (…). (Lovins 1975:6)

But bilingualism is also the source of motivation for the transformation of the 'ousted' phonological substitutions to stay productively in the language as morphonological rules, and for their generalization in their new domains. For that to be true, however, individual bilingualism must be continued by socially strong bilingualism. Somehow, we are seeing that the loss of phonological processes can lead to the birth of new morphological translation resources. We are therefore not talking about "phonologically unmotivated changes" (cf. Hualde 1993) or "essentially arbitrary rules" (Hualde 2000:348), but about changes that have a substantially different kind of motivation, which lies out of phonology proper. In other words, we are simply not talking about phonology any more. Of course the sociocultural conditions present now in the ACBC, where Basque is officially supported and ideologically prestigious among its *active* speakers, are essential for this second stage to develop, and they should be carefully studied.

---

21. Cf. Calabrese's (this volume) Introduction on the two possible sources of loanwords.

References

Browman, Catherine P. & Louis Goldstein. 1990. "Tiers in articulatory phonology". *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech* ed. by John Kingston & Mary E. Beckman, 341–376. Cambridge: Cambridge University Press.

Browman, Catherine P. & Louis Goldstein. 1992. "Articulatory phonology: an overview". *Phonetica* 49.155–180.

Busà, M. Grazia. 2007. "Coarticulatory nasalization and phonological developments: data from Italian and English nasal-fricative sequences". *Experimental Approaches to Phonology* ed. by M.J. Solé, P.S. Beddor & M. Ohala,155–174. Oxford: Oxford University Press.

Cano, Rafael. 2004. "Cambios en la fonología del español durante los siglos XVI y XVII". *Historia de la Lengua Española* ed. by Rafael Cano Aguilar, 825–857. Barcelona: Ariel.

Churma, Donald G. 1984. "Impossible nativizations as phonological evidence and the explanation of constraints of phonological structure". *Journal of Linguistics* 30.223–227.

Donegan, Patricia. 1993. "On the phonetic basis of phonological change". *Historical Linguistics. Problems and Perspectives* ed. by Charles Jones, 98–130. London: Longman.

Donegan, Patricia. 1995. "The innateness of phonemic perception". *Proceedings of the 24th Western Conference on Linguistics* ed. by Vida Samiian & Jeanette Schaeffer, vol. 7,59–69. Fresno, CA: Dept. of Linguistics, California State University.

Donegan, Patricia Jane & David Stampe. 1979. "The Study of Natural Phonology". *Current Approaches to Phonological Theory* ed. by Daniel A. Dinnsen, 126–173. Bloomington: Indiana University Press.

Dressler, Wolfgang U. 1985. *Morphonology*. Ann Arbor: Karoma.

Haase, Martin. 1992. *Sprachkontakt und Sprachwandel im Baskenland. Die Einflüsse des Gaskognischen und Französischen auf das Baskische*. Hamburg: Helmut Buske Verlag.

Hualde, José Ignacio. 1993. "Phonologically unmotivated changes in language contact: Spanish borrowings in Basque". *Folia Linguistica* 27:1–2.1–25.

Hualde, José Ignacio. 2000. Patterns of correspondence in the adaptation of Spanish borrowings in Basque". *General Session and Parasession on Loan Word Phenomena: Proceedings of Twenty-Fifth Annual Meeting of the Berkeley Linguistics Society*, 12–15 February 1999 ed. by Steve S. Chang, Lily Liaw & Josef Ruppenhofer, 348–358. Berkeley: Berkeley Linguistic Society.

Hualde, José Ignacio & Jon Ortiz de Urbina, eds. 2003. *A Grammar of Basque*. Berlin: Mouton de Gruyter.

Hurch, Bernhard. 1988a. "Phonetics and phonology or phonology and Phonetics". *ZPSK* 41:3.346–352.

Hurch, Bernhard. 1988b. "On the notion 'latent process'". *Certamen Phonologicum* ed. by Pier M. Bertinetto & Michele Loporcaro, vol. I, 103–115. Torino: Rosenberg & Sellier.

Jauregi, Oroitz & Miren Lourdes Oñederra. 2008. "Africación de sibilante tras consonante sonante". *Language Design. New Trends in Experimental Phonetics* ed. by Antonio Pamies & Elizabeth Melguizo, 135–140. Granada: Método Ediciones.

Kruszewski, Nikolai. 1978. "On sound alternation". *Readings in Historical Phonology*: *Chapters in the Theory of Sound Change* ed. by Philip Baldi & Ronald N. Werth, 64–91. The Pennsylvania State University Press. (first published as *Über die Lautabwechslungen*, Kazan, 1881).

Lovins, Julie B. 1975. *Loanwords and the Phonological Structure of Japanese*. Bloomington, Indiana: IU Linguistics Club.

Major, Roy C. 1977. "Phonological differentiation of a bilingual child". *Papers in Psycholinguistics and Sociolinguistics* ed. by Olga K. Garnica, 88–122. Ohio: Ohio State University.

Michelena, Luis. 1977. *Fonética Histórica Vasca*. San Sebastián: Seminario Julio de Urquijo de la Excma. Diputación de Guipúzcoa.

Michelena, Luis. 1995. "The Latin and Romance element in Basque". *Towards a History of the Basque Language* ed. by José Ignacio Hualde, Joseba A. Lakarra & R.L. Trask, 137–169. Amsterdam & Philadelphia: John Benjamins.

Ohala, John. 1981. "Speech timing as a tool in phonology". *Phonetica* 38.204–212.

Oñederra, Miren Lourdes. 2002. "Palatalization revisited: bustidura adierazgarriaren estatusaz ohar batzuk [Palatalization revisited: notes on the status of expressive palatalization]". *Erramu Boneta*: *Festschrift for Rudolf P. G. de Rijk* ed. by X. Artiagoitia, P. Goenaga & J. Lakarra, 499–509. (= *ASJU*, Anejo XLIV). Bilbao: Universidad del País Vasco/Euskal Herriko Unibertsitatea.

Oñederra, Miren Lourdes (in prep). El patrón sonoro de la lengua vasca.

Pensado, Carmen. 1999. "Sobre la interpretación de lo inexistente: los tipos silábicos inexistentes en la fonología del español". *Panorama de la Fonología Española Actual* ed. by Juana Gil Fernández, 475–483. Madrid: Arco Libros. (first published in 1985 as "On the interpretation of the non-existent: non-occurring syllable types in Spanish phonology". *Folia Linguistica* XIX.313–319).

Picard, Marc & Janet Nicol. 1982. "Vers un modèle concret de la phonologie des emprunts". *CJL/RCL* 27:2,156–169.

Stampe, David. 1969. "The acquisition of phonetic representation". *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society* ed. by Robert I. Binnick, Alice Davison, Georgia Green & Jerry L. Morgan, 443–454. Chicago: University of Chicago Department of Linguistics. (= Stampe 1979. vii-xxv).

Stampe, David. 1979. *A Dissertation on Natural Phonology*. New York: Garland Publishing Co.

Trask, R.L. 1997. *The History of Basque*. London: Routledge.

Wurzel, Wolfgang U. 1980. "Ways of morphologizing phonological rules". *Historical Morphology* ed. by Jacek Fisiak, 443–462. The Hague: Mouton.

Nondistinctive features in loanword adaptation

The unimportance of English aspiration in Mandarin Chinese phoneme categorization

Carole Paradis & Antoine Tremblay
Laval University/University of Alberta

Based on a corpus of 500 stops included in 371 borrowing forms from English in Mandarin Chinese (MC), we show that English stop aspiration, which is agreed to be phonetic, does not influence phoneme categorization in MC, despite the fact that MC has phonemic aspirated stops. Thus even if their mother tongue predisposes MC speakers to distinguish aspirated from unaspirated stops, they do not rely on aspiration in English to determine phoneme categorization in MC. Both aspirated and unaspirated voiceless stops of English systematically yield an aspirated stop in MC, whereas English voiced stops, which are disallowed in MC, systematically yield a voiceless unaspirated stop. These facts disfavor the perceptual stance in loanword adaptation and lend support to the phonological one.

1. Introduction

Stop aspiration in English is agreed to be phonetic since it is predictable, and thus nondistinctive. It occurs when the following conditions are met (Rogers 2000):

(1)   Conditions of stop aspiration in English[1]

    a.   Aspiration applies to a voiceless stop.
    b.   This voiceless stop must be part of an onset.
    c.   It must not be preceded by /s/.
    d.   It must be followed by a nucleus bearing primary or secondary stress.

This article aims to show that English stop aspiration, because it is phonetic in English, constitutes irrelevant information in the adaptation of loanwords. Here we will be concerned with English loanwords in Mandarin Chinese (MC). The case of

---

[1]   "Voiceless stops are aspirated at the beginning of a stressed syllable : … However, after a syllable initial /s/ or at the beginning of an unstressed syllable, voiceless stops are not aspirated …" Rogers (2000:50). Also see Odden (2005:46).

MC is particularly interesting given the fact that (i) MC has aspirated stops which, contrary to English, are phonological (categorical) as opposed to phonetic (noncategorical), and (ii) MC has no categorical voiced stops. This is shown in Table 1.

Table 1. MC's consonant inventory (based on Duanmu 2002:25–26)

	Labials	Dentals	Retroflexes	Palatals	Velars
Stops	p	t			k
	pʰ	tʰ			kʰ
Affricates		ts	tʂ		
		tsʰ	tʂʰ		
Fricatives	f	s	ʂ		x
Nasals	m	n			ŋ
Liquids		l	ɻ		
Glides	w			j	
				ɥ	

As addressed at length in LaCharité & Paradis (2005), there are two opposing views regarding the kind of information that is relevant to sound adaptation in loanwords: the perceptual stance, which maintains that crucial information to loanword adaptation is phonetic (see, for instance, Silverman 1992; Yip 1993; Kenstowicz 2003; Peperkamp & Dupoux 2002, 2003; Hsieh, Kenstowicz, & Mou, this volume), and the phonological stance, according to which exclusively distinctive information is relevant to loanword adaptation (see, for example, Hyman 1970; Danesi 1985; Paradis & LaCharité 1997; Paradis & Prunet 2000; Ulrich 1997; Jacobs & Gussenhoven 2000; Davis & Kang 2003). Regarding English stop aspiration, these two stances make the following opposing predictions:

(2)   Opposing predictions

a.   The perceptual stance: Since phonetic details matter in the adaptation of loanwords, aspirated stops in English should systematically yield aspirated stops in MC, whereas English unaspirated stops should remain unaspirated in MC.

b.   The phonological stance: Since phonetic details do not matter in the adaptation of loanwords, there should not be a direct correlation between English aspirated/unaspirated voiceless stops and MC aspirated/unaspirated ones.

Our results, which will be provided in the third section along with a description of our corpus and statistics, support the phonological view, not the perceptual one. It appears that the distinction between voiced and voiceless stops in English is preserved in MC with another laryngeal feature: aspiration. Foreign voiced stops are adapted as unaspirated voiceless stops (e.g., English *Boeing* [ˈboɪŋ] > MC [pəin]), whereas foreign voiceless ones are systematically adapted as aspirated

ones (e.g., English *pizza* [ˈpʰitsə] > MC [pʰitsa]/[pʰisa]). Quite significant is the fact that English voiceless stops yield aspirated stops in MC even when they are not aspirated in English (e.g., English *hippies* [ˈhɪpiz] *[ˈhɪpʰiz] > MC [sipʰiʂ]).

LaCharité & Paradis (2005) show that naive phonetic approximation, as opposed to categorical (phonological) adaptation, plays a very minor role in Project CoPho's large database of loanwords (more than 50,000 malformations, that is foreign phonemes and structures, from several corpora).[2] Phonetic details are basically relevant only in what they call intentional phonetic approximation, that is importations (nonadaptations) of foreign sounds and structures.

These conclusions are not in line with those reached by Kang (2003), who studied English postvocalic word-final stops in Korean. From a list of loanwords compiled by the National Academy of the Korean Language, which contains 5,000 English words and phrases gathered from newspapers and magazines, Kang (2003:220) shows that the insertion of nondistinctive Korean vowels in loanwords from English is more likely "(a) when the [English] pre-final vowel is tense rather than lax, (b) when the [English] final stop is voiced rather than voiceless and (c) when the [English] final stop is coronal rather than non-coronal and when the stop is labial rather that dorsal". Given that English final stops are more often than not released after tense rather than lax vowels, nondistinctive vowel insertion in Korean is believed to yield "good perceptual approximation to [English] stop release" (Kang, 2003:220). Silverman (1992) makes similar claims concerning English stop aspiration in Cantonese Chinese. He reports, for instance, that English *tie* [tʰaj] yields [tʰaj] in Cantonese Chinese, whereas English *stick* [stɪk], whose stop is unaspirated, yields [sɪtɪk]. While the first example yields a result that is consistent with either view — with the perceptual stance, because the source is aspirated and, with the phonological stance, because a voiceless stop is systematically adapted as an aspirated one in Cantonese Chinese — the second example is less in accordance with the phonological view.

In the present paper, we provide and discuss evidence coming from MC, which lends support to the phonological stance (i.e., noncontrastive phonetic details are not important in sound adaptation).

2.   Methodology

Our corpus of English loanwords in MC includes 77 borrowings containing voiced and voiceless stops. All the stops of our study are contained in phonological borrowings, not semantic ones. What we mean by 'semantic borrowing' is best clarified by

---

2.   Project CoPho, which is supervised by Carole Paradis at Laval University, is concerned with the role of constraints (Co) in phonology (Pho).

way of example. English *shampoo* /ʃæmˈpu/ is adapted in MC as /sjaŋ–pə/, where /sjaŋ/ means 'perfumed' and /pə/ 'wave', meaning literally 'perfumed wave'. Here both syllables in the English word *shampoo* are adapted by analogy to a phonologically similar word in an MC construction that is semantically related to the English meaning. Semantic borrowings are common in Chinese but since they are adapted by analogy (either true or false) to a Chinese word, they are not relevant to phonology and thus to our study.

The borrowings were collected from various written and oral sources and are all attested in the *Xiandai Hanyu Cidian* [Dictionary of Modern Chinese], the *Ci Hai* [Shang Hai Dictionary], Au-Yeung (1997), Liu (1995), and/or Miao (2005). All the borrowings were introduced in MC after 1919, except for three. Most are much more recent. In 2003, we solicited their pronunciations from five MC native speaking informants: Two are from Beijing (a 47-year-old woman and a 33-year-old man), one is from Tianjin (a 38-year-old woman), and two are from the province of Hubei (a 26-year-old woman and a 34-year-old man). Of the 77 borrowings, the informants pronounced the borrowings that were known to them. This yielded 371 borrowing forms (i.e., the concrete realizations of the borrowings) containing 500 stops (363 of them voiceless and 137 voiced). We elicited borrowings that refer to concrete objects through picture-naming tasks; definitions, paraphrases, and fill-in-the-blank sentences were used to elicit abstract borrowings. The sessions were conducted in MC by the second author, who is fluent in this language. All the forms were tape-recorded and transcribed in IPA, in what Duanmu (2002) considers phonological form, as much as possible.[3] Transcriptions were checked by two phoneticians/phonologists whose mother tongue is MC before being computerized. As for our English transcriptions, they are adapted from those found in the *Longman Dictionary of Contemporary English*. For purpose of uniformity with the rest of Project CoPho's database of loanwords, we transcribed the Longman symbols in the following way: ʌ > ə, aɪ > aj , ɜɪ > ər, eɪ> e , oʊ > o, ɹ > r, and ɔ > a.

3.   The Mandarin Chinese language

Turning now to the relevant aspects of the target language's phonology, MC has five vowels : /i, y, u, ə, a/. There might be a sixth one, /ɻ/ (transcribed as [r] for convenience in Duanmu 2002), but its phonological status is uncertain according to Duanmu (2002:37–42). MC's maximal syllable is C(G)V(V/C), as shown in Figure 1.

---

3.   For example, [jii] was transcribed /i/ in the corpus because the phonological form is /i/. Here, vowel spreading in the onset and vowel lengthening are both predictable.
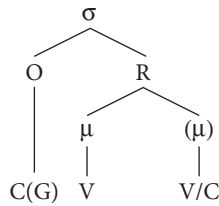
Figure 1. Maximal syllable in MC (adapted from Duanmu 2002:45)

The glide (G) in Figure 1 constitutes a complex consonant with the preceding segment. Rhymes are usually bimoraic, that is, they are composed of either a long vowel or a short vowel and a coda. Any consonant, except /ŋ/, can appear in an onset but only the nasals /n/ and /ŋ/, the retroflex /ɻ/, and the glides /j/ and /w/ occur in coda position. Bimoraic syllables are deemed "heavy" whereas syllables with a short vowel are designated as "light." On the metrical side, words in MC have a strong tendency to be disyllabic (see Malischewski 1987; Good 1996; Wang 2004; Hu 2004), although words of more than three syllables are attested in rare cases.

4. Results and discussion

4.1 Voiceless English stops

As shown in Figure 2 and Table 2, in our corpus of English loanwords in MC, we find 363 voiceless stops, 83.2% of which are aspirated in MC.
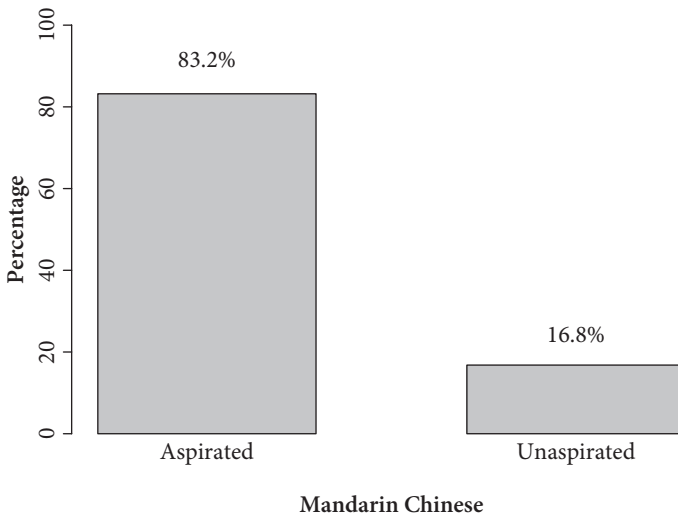
Figure 2. English voiceless stops /p, t, k/ in MC

Table 2. Aspirated and unaspirated English /p, t, k/ segments in Mandarin Chinese

	363 cases of English /p, t, k/
Aspirated in MC	83.2 % (302/363)
Unaspirated in MC	16.8% (61/363)

A chi-square test revealed that the two values are significantly different ($\chi^2 = 160$, $df = 1$, $p < .001$).[4]

In our corpus, 160/363 English voiceless stops appear in a context of aspiration and 203/363 in a context of nonaspiration. Table 3 shows how the English aspirated segments [p^h, t^h, k^h] were adapted in MC.

Table 3. Cases of English voiceless aspirated [p^h, t^h, k^h] stops in MC

	[p^h]	[t^h]	[k^h]	Total
Aspirated in MC	21/32	47/48	74/80	142/160
	65.6%	97.9%	92.5%	88.8%
Unaspirated in MC	11/32	1/48	6/80	18/160
	34.4%	2.1%	7.5%	11.2%

Note: [p^h]: $\chi^2 = 3$, $df = 1$, $p = .077$; [t^h]: $\chi^2 = 44$, $df = 1$, $p < .001$; [k^h]: $\chi^2 = 57$, $df = 1$, $p < .001$.

Table 3 shows a significant difference between aspirated and unaspirated segments in MC. That is, English aspirated stops are adapted as aspirated stops in MC significantly more often than they are adapted as unaspirated stops. Table 4 shows the same for English [p, t, k].

Table 4. Cases of English voiceless unaspirated [p, t, k] stops in MC

	[p]	[t]	[k]	Total
Aspirated in MC	38/44	31/58	91/101	160/203
	86.4%	53.4%	90.1%	78.8%
Unaspirated in MC	6/44	27/58	10/101	43/203
	13.6%	46.6%	9.9%	21.2%

Note: [p]: $\chi^2 = 23$, $df = 1$, $p < .001$; [t]: $\chi^2 = 0.3$, $df = 1$, $p > .05$; [k]: $\chi^2 = 65$, $df = 1$, $p < .001$.

---

4.   Statistics were computed using *R*, a statistical analysis package, which can be downloaded from www.R-project.org.

Specifically, Table 4 shows that unaspirated English stops are also adapted as aspirated stops in MC significantly more often than they are adapted as unaspirated stops. As predicted by the phonological stance in (2b), voiceless stops, whether they are aspirated in English or not, yield an aspirated voiceless stop in MC (Aspirated in English: $\chi^2 = 96.1$, $df = 1$, $p < .001$; Unaspirated in English: $\chi^2 = 67.4$, $df = 1$, $p < .001$). This is illustrated in Figures 3 and 4.
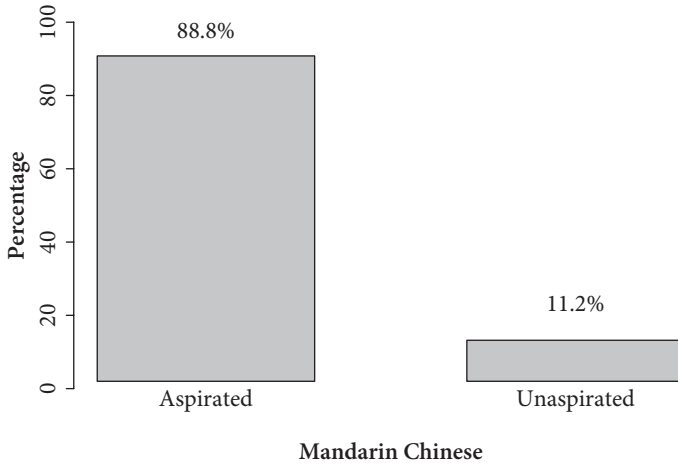
Figure 3. Voiceless aspirated [pʰ, tʰ, kʰ] stops in English yielding aspirated stops in MC
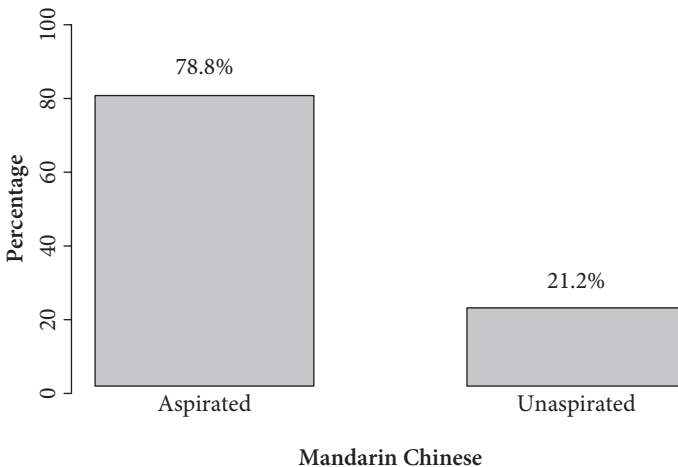
Figure 4. Voiceless unaspirated [p, t, k] stops in English yielding aspirated stops in MC

Figure 4 strongly suggest that aspiration — a phonetic detail in English — does not influence categorization (i.e., phonological adaptation). Crucially, 160/203 (78.8%) cases of English voiceless stops in a nonaspiration context yield an aspirated stop in MC. Examples of unaspirated stops in English yielding aspirated stops in MC are presented in Table 5.

Table 5. Examples of unaspirated stops in English yielding aspirated stops in MC

English	IPA	MC
a. ampere	/ˈæmpeɹ/	/an pʰəj/
b. chocolate	/ˈtʃaklɪt/	/tsjaw kʰ ə li/
c. Intel	/ˈintɛl/	/in tʰə ər/
d. internet	/ˈint əɹnɛt/	/in tʰə waŋ/
e. jacket	/ˈdʒækɪt/	/tsja kʰə/
f. microphone	/ˈmajkɹ əfon/	/maj kʰə fəŋ/
g. Olympic	/ əˈlɪmpɪk/	/aw lin pʰi kʰə/
h. opium	/ˈopjəm/	/ja pʰjan/
i. poker	/ˈpokəɹ/	/pʰu kʰə/
j. quark	/kwaɹk/	/kʰwa kʰə/
k. totem	/ˈtotəm/	/tʰu tʰəŋ/
l. tank	/tæŋk/	/tʰan kʰə/
m. volt	/volt/	/fu tʰə/
n. watt	/wat/	/wa tʰə/

Note that a voiceless stop in word-final position, which is never aspirated in English (since there is obviously no following stressed nucleus), also yields an aspirated voiceless stop in MC when a vowel is epenthesized after it (e.g., /aw lin pʰi kʰə/ < *Olympic*, /fu tʰə/ < *volt*, /kʰwa kʰə/ < *quark*, and /tʰan kʰə/ < *tank*). This clearly supports the hypothesis that a foreign voiceless stop is automatically encoded (i.e., categorized) as an aspirated voiceless stop in MC and that its phonetic realization in the donor language does not determine its categorization in the borrowing language.

As for English voiceless stops that are realized as unaspirated stops in MC, they might be importations (i.e., nonadaptations), possibly imperfect ones in some cases. Imperfect importations do not necessarily display all the phonetic details they include in the donor language (in this case, aspiration). Whatever the reasons, only a few cases are at stake here.

4.2   Voiced English stops

The reason English voiceless stops are categorized as aspirated voiceless stops in MC is because MC does not have voiced stops. English voiced stops are encoded as voiceless unaspirated stops in MC (89.8% of the cases; $\chi^2 = 86.7$, $df = 1$, $p < .001$), as shown in Table 6.

Table 6. English voiced stops yielding unaspirated voiceless stops in MC

	/b/	/d/	/g/	Total
Aspirated in MC	1/41	12/76	1/20	14/137
	2.4%	15.8%	5.0%	10.2%
Unaspirated in MC	**40/41**	**64/76**	**19/20**	**123/137**
	97.6%	**84.2%**	**95.0%**	**89.8%**

Note: /b/: $\chi^2 = 37.1$, $df = 1$, $p < .001$; /d/: $\chi^2 = 35.6$, $df = 1$, $p < .001$; /g/: $\chi^2 = 16.2$, $df = 1$, $p < .001$.

The fact that English voiced stops are adapted as unaspirated stops in MC significantly more often than as aspirated stops is further illustrated in Figure 5.
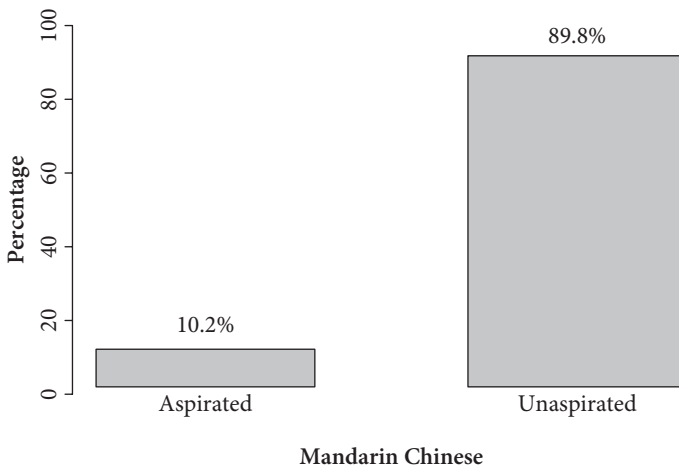
Figure 5. English voiced stops /b, d, g/ yielding unaspirated voiceless stops in MC

Examples of English voiced stops yielding voiceless unaspirated stops in MC are provided in Table 7.

Table 7. Examples of English voiced stops yielding voiceless unaspirated stops in MC

English	IPA	MC
a. bandage	/ˈbændɪdʒ/	/paŋ ti/
b. bar	/baɹ/	/tsjəw pa/[*]
c. carbine	/ˈkaɹbajn/	/kʰa pin tsʰjaŋ/[**]
d. golf	/galf/	/kaw ər fu/
e. guitar	/gɪˈtaɹ/	/tsi tʰa/
f. radar	/ˈɹedaɹ/	/ləj ta/
g. sardine	/saɹˈdin/	/ʂa tiŋ y/[†]

Notes: [*] /tsjəw/ means "alcohol". [**] /tsʰjaŋ/ means "gun". [†] /y/ means "fish"

English voiced stops that yield an aspirated stop in MC (14/137 cases, 10.2%) stem mostly from two borrowings: English *model* /ˈmadəl/ and *mandolin* /ˈmændəlɪn/ that are adapted as MC /mə tʰə/ and /man tʰwə lin/ respectively. These forms are possibly influenced by a nonphonological factor, such as analogy, but we were not able to identify it. In any case, although the voiced stop in these forms is not adapted according to the regular phonological pattern, that is /t/, neither can resulting /tʰ/ be accounted for by the phonetic stance. Phonetically, English voiced stops are more closely related to unaspirated MC stops than to aspirated ones in terms of their VOT values, whether the English voiced stop is pronounced with voicing lag (positive VOT values) or voicing lead (negative VOT values). This is illustrated in Table 8.[5]

By way of example, English /d/ has a VOT of 5 msecs when pronounced with voicing lead and –102 msecs when uttered with voicing lag. On the other hand, MC /t/ and /tʰ/ have a VOT value of 9 and 74 msecs respectively. This means that English /d/ produced with voicing lead or lag is closer to MC /t/ (difference of –4 and –106 msecs) than it is to MC /tʰ/ (difference of –69 and –176 msecs).

In other words, if we consider VOT norms, the phonetic stance as well as the phonological one predicts that English such as /d/ in *model* /ˈmadəl /, should be adapted as MC voiceless unaspirated stops (e.g., /t/) since they are closer in terms of their VOT values.

---

5. VOT values for English voiced and voiceless stops are adapted from Lisker & Abramson (1964:394-395). VOT values for MC stops are taken from Zhao & Meng (1997:77).

Table 8. Distance between English voiced stops and MC unaspirated and aspirated stops in milliseconds (msecs)

		English VOT Norms (in msecs)		MC VOT Norms (in msecs)		Differences in VOT norms (in msecs)
b	Lead	1	p	13		−12
		1	p^h	67		−66
b	Lag	−101	p	13		−114
		−101	p^h	67		−168
d	Lead	5	t	9		−4
		5	t^h	74		−69
d	Lag	−102	t	9		−111
		−102	t^h	74		−176
g	Lead	21	k	21		0
		21	k^h	75		−54
g	Lag	−88	k	21		−109
		−88	k^h	75		−163

5. Conclusion

In this article, we have endeavored to show that English stop aspiration, which is phonetic, does not influence phoneme categorization in MC, in spite of the fact that MC has phonemic aspirated stops. In other words, despite the fact that their mother tongue predisposes MC speakers to distinguish aspirated from unaspirated stops, they do not rely on aspiration/nonaspiration in English to determine phoneme categorization in MC. Both aspirated and unaspirated voiceless stops of English systematically yield an aspirated stop in MC, whereas English voiced stops systematically yield a voiceless unaspirated stop in MC.

Kim (this volume) obtained identical results in Korean concerning the adaptation of English stops, that is, English voiceless stops are systematically aspirated in korean whereas voiced stops yield unaspirated voiceless stops. Hindi also seems to pattern in a similar way. The corpus of English loanwords in Hindi that Project CoPho is currently assembling also shows that the aspiration of stops in English has no impact on the way these are adpated in Hindi.Voiceless stops are systematically adapted as non-aspirated voiceless stops and voiced stops as unaspirated voiced stops despite the fact that aspiration is distinctive in Hindi.

These facts add to the numerous arguments provided, for instance, in Jacobs & Gussenhoven (2000), Paradis & Prunet (2000), Paradis & LaCharité

(2001), Davis & Kang (2003), LaCharité & Paradis (2005), and more recently Paradis (2006) against the perceptual stance and lend further support to the phonological one.

As shown by Calabrese and Wetzels in their introduction to this volume, the contributions gathered here support evidence for both the perceptual stance (see, e.g., Hsieh et al.) and the phonological one (our contribution), as well as for intermediate views (see, e.g., Kim). Analogously to the blind men touching different parts of the elephant, and falsely inducing the presence of a snake, a rope, a wall, etc., the two stances presented and advocated for in this volume is necessarily incomplete at this stage of our knowledge and may well turn out to be part of the same system, albeit with different *raisons d'être*.

Can all these views be reconciled? We believe that they can and that differences of results are often attributable to differences of conceptualization, terminology and methodology. As pointed out in Paradis & LaCharité's (to appear) article, the importance of methodology should not be underestimated. As in any area of science, different methodologies may lead to different results. This is why it is so important to be explicit about the methodology used in the study of loanword adaptation and to work with statistically based corpora/data. Perhaps apparently contradictory results can be reconciled when issues of methodology are taken into account. As shown in Paradis & LaCharité (2008), phonetic approximation exists in the CoPho loanword database, but it is weak overall. Things are often not what they first seem when more factors, such as the prestige of the L2, negative feelings towards the L2, normalization, analogy, either false or real, the distinction between naive and intentional phonetic approximation (real phonetic approximation vs. importation of foreign phonemes/structures), native processes, hypercorrection, are taken into account.

Acknowledgements

References

Au-Yeung, Annie. 1997. *Lexique des mots nouveaux de la langue chinoise.* Paris: Éditions You Feng.
Ci Hai. 1979. Shang Hai: Shang Hai Ci She Chu Ban She.

Danesi, Marcel. 1985. *Loanwords and phonological methodology*. Ville LaSalle, Que.: Didier.

Davis, Stuart & Hyunsook Kang. 2003. "On two phonetic mismatches in the borrowings of English words into Korean". Paper presented at the Midcontinental Workshop on Phonology, University of Illinois, Urbana, October 2003.

Duanmu, San. 2002. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Good, Robert. 1996. "The automatic identification of English transliterated loanwords in Chinese". *Studies in the Linguistic Sciences* 26.133–142.

Hu, Peter. 2004. "Adapting English into Chinese". *English Today* 20.34–39.

Hyman, Larry. 1970. "The role of borrowings in the justification of phonological grammars". *Studies in African Linguistics* 1.1–48.

Jacobs, Haike & Carlos Gussenhoven. 2000. "Loan phonology: Perception, salience, the lexicon and OT". *Optimality Theory: Phonology, syntax, and acquisition* ed. by Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer, 193–210. Oxford: Oxford University Press.

Kang, Yoonjung. 2003. "Perceptual similarity in loanword adaptation: adaptation of English postvocalic word-final stops to Korean". *Phonology* 20.219–273.

Kenstowicz, Michael. 2003. "The role of perception in loanword phonology: A review of *Les emprunts linguistiques d'origine européenne en Fon*". *Studies in African Linguistics*, 32.95–112.

LaCharité, Darlene & Carole Paradis. 2005. "Category preservation and proximity versus phonetic approximation in loanword adaptation". *Linguistic Inquiry* 36.223–258.

Lisker, Leigh & Arthur S. Abramson. 1964. "A cross-language study of voicing in initial stops: Acoustical measurements". *Word* 20.384–422.

Liu, Lydia H. 1995. *Translingual practice. Literature, national culture and translated modernity. China 1900–1937*. Stanford, Calif.: Stanford University Press.

Longman Dictionary of Contemporary English. 2000. Harlow: Pearson Education.

Malischewski, E. 1987. "Kai Fang: Loan words to the Middle Kingdom". *English Today* 12.40–41.

Miao, Ruiquin. 2005. *Loanword adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors*. Doctoral dissertation, Stony Brook University, New York.

Odden, David. 2005. *Introducing Phonology*. Cambridge: Cambridge University Press.

Paradis, Carole & Darlene LaCharité To appear. "Loanword Adaptation: From Lessons Learned to Findings". *Handbook of Phonological Theory* ed. by John Goldsmith, 2nd edition, 45 p. Cambridge: Blackwell.

Paradis, Carole & Darlene LaCharité. 2008. "Apparent Phonetic Approximation: English Loanwords in Old Quebec French". *Journal of Linguistics* 44:87–128.

Paradis, Carole. 2006. "The unnatural /Ċu/ (< foreign /Cy/) sequence in Russian loanwords: A problem for the perceptual view". *Special issues in loanword Phonology* ed. by Michael Kenstowicz & Christian Uffman, *Lingua* 116.976–995.

Paradis, Carole & Darlene LaCharité. 1997. "Preservation and minimality in loanword adaptation". *Journal of Linguistics* 33.379–430.

Paradis, Carole & Darlene LaCharité. 2001. "Guttural deletion in loanwords". *Phonology* 18.255–300.

Paradis, Carole & Jean-François Prunet. 2000. "Nasal vowels as two segments: Evidence from borrowings". *Language* 76.324–357.

Peperkamp, Sharon & Emmanuel Dupoux. 2002. "Loanword adaptations: Three problems for phonology (and a psycholinguistic solution)". Paper presented at the *North American Phonology Conference (Naphc)*, Concordia University, Montreal, April 27 2002.

Peperkamp, Sharon & Emmanuel Dupoux. 2003. "Reinterpreting loanword adaptations: The role of perception". *Proceedings of the 15th International Congress of Phonetic Sciences*, 367–370. Barcelona : Causal Productions.

R Development Core Team. 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rogers, Henry. 2000. *The Sounds of language : An introduction to phonetics*. Toronto: Pearson Education Ltd.

Silverman, Daniel. 1992. "Multiple scansions in loanword phonology: Evidence from Cantonese". *Phonology* 9.289–328.

Ulrich, Charles. 1997. "Loanword adaptation in Lama: Testing the TCRS model". *Canadian Journal of Linguistics* 42.415–463.

Wang, Hongyuan. 2004. "Loans from the Middle Kingdom revisited". *English Today* 20.47–49.

Xiandai Hanyu Cidian [Dictionary of Modern Chinese]. 1998. Bei Jing: Shang Wu Yin Shu Guan.

Yip, Moira. 1993. "Cantonese loanword phonology and Optimality Theory". *Journal of East Asian Linguistics* 2.261–291.

Zhao, Jingming & Ziming Meng. 1997. *Yuyin Yanjiu yu Duiwai Hanyu Jiaoxue* [Research in Phonetics and Teaching Mandarin Chinese as a Second Language]. Beijing: Beijing yuyan wenhua daxue chubanshe.

Gemination in English loans in American varieties of Italian*

Lori Repetti
SUNY, Stony Brook

Why do geminate consonants frequently appear in borrowed words when the foreign form does not contain a geminate? In this paper I review previous approaches to this problem, and suggest that they are insufficient in accounting for consonant length contrasts in English loan words in North American varieties of Italian. I suggest that many factors are involved in the determination of consonant length in loans, including aspects of the grammar of the borrowing language (in this case, Italian) — such as the inventory of segments, the structure of the stressed syllable, and the presence of similar native lexical items — as well as the interpretation of the morphological structure and phonetic details of the foreign word.

1. Introduction

Non-etymological geminates often appear in the adapted form of loan words, and are attested in Japanese, Finnish, Kannada, Maltese Arabic, Hungarian, and Italian, including North American varieties of Italian (henceforth "American-Italian") (1), as well as many other languages.

(1)	English	American-Italian
a.	*coal*	[ˈkolle]
b.	*gingerale*	[ˈdʒindʒaˈrella]
c.	*brush*	[ˈbroʃʃa]
d.	*bushel*	[ˈbuʃʃolo]
e.	*creek*	[ˈkrikka]
f.	*fell*	[ˈfelli]
g.	*tape*	[ˈteppa]
h.	*team*	[ˈtimme]

The phenomenon whereby a singleton consonant in the loaning language is adapted as a geminate consonant in the borrowing language is very common cross-linguistically, and has vexed phonologists for some time. In this paper, I will show that a wide variety of factors — including phonetic, phonological, morphological, and lexical considerations — may come into play in determining which, if any, consonants will lengthen in the integration of foreign loans. I will illustrate this approach using data from American-Italian, which is the variety of Italian spoken by Italian immigrants to North America whose native language is/was an Italian dialect. (Data are from various published sources referenced in the Bibliography and from field research.)

Analyses of gemination in loans abound in the literature. Gemination of the consonant following the stressed vowel has been attributed to syllable structure, whereby borrowers try to preserve the syllable structure of the foreign form, and specifically the moraicity of final consonants, through gemination (Katayama 1998). Metrical requirements have also been invoked: if the stressed syllable in the borrowing language must be bimoraic, gemination is a means of satisfying this requirement (Repetti 1993). It has also been claimed that morpho-phonological alignment constraints are at work: the foreign noun is identified as a stem, the stem must be aligned with a syllable, and gemination is the means by which this require-ment is met (Shinohara 2003; Repetti 2003, 2006). Finally, it has been proposed that borrowers interpret fine acoustic details of vowel and consonant length in terms of their own phonological system, rendering the consonant following a (phoneti-cally) short vowel as long, and the consonant following a (phonetically) long vowel as short (Abraham 2004; Peperkamp & Dupoux 2003).

While these approaches can account for some of the data in (1), the problem is that not all English loan words in American-Italian undergo gemination (2). Fur-thermore, those that do and those that do not, don't seem to form natural classes. (Compare the forms in (1) with the data in (2).)

(2)	English	American-Italian
a.	*bowl*	['bolo]
b.	*wholesale*	[ol'sele]
c.	*bruise*	['brusa]
d.	*people*	['pipoli]
e.	*strike*	['strajko]
f.	*fellow*	['falo]
g.	*paper*	['pepa]
h.	*steam*	['stima]

We will see that the American-Italian data are not consistent with any *one* of the analyses mentioned above. Instead, a combination of factors is needed to determine consonant length in loan words: lexical considerations, morpho-phonological constraints, and perceptual factors. In particular, I show how the following factors

play a role in determining consonant length in loans: the presence of lexical items in the borrowing language with a similar phonological structure and a compatible meaning (§2), facts about the phonemic inventory and syllable structure of the borrowing language (§3–§4), morpho-phonological alignment constraints (§5), the interpretation of acoustic details of the foreign words (§6), and a universal principle regarding sonorant geminates (§6). I show how, within the framework of Optimality Theory (OT) (Prince & Smolensky 1993; McCarthy & Prince 1993), we can account for the American-Italian data by using ranked constraints. (See also Friesner in this volume for a discussion of various social and grammatical factors affecting loanword nativization in Romanian.)[1]

2. Similar native lexical items

If the foreign word is similar in phonological form to a word in the borrowing language, and the two words have compatible meanings, the native word (along with its consonant length) is used.

(3) <u>English</u>   <u>American-Italian</u>
    coal   *colle* (standard Italian 'hill')
    furniture   *fornitura* (standard Italian 'supply')

If the English word ends in a series of segments identified as an Italian suffix, that suffix is used along with its lexically determined consonant length.

(4) *basket*   [bas'ketto]   (diminutive suffix: -étto/a)
    machine   [maʃ'ʃina]   (diminutive suffix: -íno/a)
    ginger ale   [dʒindʒa'rella]   (diminutive suffix: -éllo/a)
    coupon   [ku'pone]   (augmentative suffix: -óne)
    bricklayer   [brikkaʎ'ʎere]   (agentive suffix: iére)
    contractor   [kontrat'tore]   (agentive suffix: óre)

These considerations outweigh any other phonetic, phonological, or morphological considerations that affect consonant length.

---

1.   There are a few aspects of the data that I will not address in this paper. (i) A final vowel is added to consonant-final English words. The quality of the final vowel is determined by morphological considerations that are not directly relevant for the question of gemination. I will not discuss these facts in this article. (ii) The Italian mid vowels are all transcribed as tense, although their tenseness may vary. This is also irrelevant for the current purposes. (iii) The position of the stressed vowel in the adapted form is usually the same as in the etymological form (see Kenstowicz 2003 for discussion of the Max-Stress constraint). However, there are exceptions, and I will not discuss the principles that determine when stress is shifted, and which syllable it is shifted to.

3.   Segments

The length of a segment may be determined by the phoneme inventory of the borrowing language. Italian has a series of consonant that must be long in inter-vocalic position: /ts/, /dz/, /ʃ/, /ɲ/, /ʎ/ (Chierchia 1986). When an English word contains one of these sounds, it is adapted as long in intervocalic position.

(5)   *peanuts*   [pi'notstsa]
      brush   ['broʃʃa]
      Flatbush   [fla'buʃʃe]

In many varieties of Italian, including most northern varieties, the dental fricative has two realizations in intervocalic position: the geminate fricative is always voiceless intervocalically [ss], and the singleton fricative is always voiced intervocalically [z]. In other words, in some varieties, a singleton voiceless [s] is not allowed in intervocalic position, and a geminate voiced [zz] is not permitted at all. Hence, the difference between the singleton and geminate dental fricative in intervocalic position is not just one of length, but also voice.

(6)   Italian: *cassa*   ['kassa] 'case'
                 casa   ['kaza] 'home'

Not surprisingly, when an English word contains a voiceless alveolar fricative, that segment is borrowed as long in intervocalic position (7a.) And when an English word contains a voiced alveolar fricative, that segment is realized as short in intervocalic position (7b).

(7)   <u>English</u>   <u>American-Italian</u>
   a. *lease*   ['lissa]
       fussy   ['fassi]
   b. *bruise*   ['bruza]
       crazy   ['krezi]

4.   Syllable structure

The metrical structure of the borrowing language can also play an important role in determining consonant length. For example, Italian allows optimally and maximally bimoraic tonic syllables. If the stressed syllable contains a falling diphthong ([aj], [aw], [oj]), which I analyze as bimoraic, the following consonant is always short. If the consonant following the diphthong were long, the stressed syllable would contain an unacceptable trimoraic structure.

(8)   *strike*   ['strajko]/*['strajkko]
      pipe   ['pajpa]/*['pajppa]
      unemployment   [anem'plojme]/*[anem'plojmme]

If there happens to be a falling diphthong plus a consonant that we expect to be long because of the segmental considerations mentioned above in §3, the conflict is resolved in favor of the diphthong, and not the consonant.

(9)   *house*   ['hauza̠]/*['haussa]

We can illustrate this within the framework of Optimality Theory, by positing four constraints ranked in a particular order relative to each other, and crucially the familiar markedness >> faithfulness ranking common in loan word adaptations.

(10)   *3μ — no trimoraic syllables (Kager 1999)
       *VsV — no intervocalic short voiceless dental fricative
       Ident-I-O(diphthong) — no changes to diphthongs
       Ident-I-O(voice) — no changes in voicing (Kager 1999)

(11)   ranking: *3μ, *VsV, Ident-I-O(diphthong) >> Ident-I-O(voice)

(12)

backhouse	*3μ	*VsV	Ident-I-O(diph)	Ident-I-O(voice)
a. [bak.ʔkaus.sa]	*!			
b. [bak.ʔkau.sa]		*!		
c. [bak.ʔkas.sa]			*!	
d. [bak.ʔkau.za]←				*

We have just seen how markedness and faithfulness constraints interact in the loan adaptation process. In the next section we will see how morphological considerations — and specifically the identification of the stem — affect loan word adaptation and consonant length.

5.   Structure of the English word

Whether or not the consonant following the stressed vowel is geminated, may depend on the structure of the English word, and, in particular, the position of the stressed vowel and whether the word ends in a vowel or a consonant. (For more on the treatment of consonant-final lexical items in Florentine and Neapolitan, see Bafile 2002.)

5.1   English v́cv#

If the English word contains a final unstressed vowel, the consonant following the stressed vowel is not geminated. (In older/other varieties of American-Italian, English words ending in an unstressed /i/ or /o/ were pronounced with stress shifted to the

final vowel, and again no gemination, as in *fellow* [fa.'lo]. See Repetti 2003, 2006 for discussion.)

(13) *fellow*   ['falo]
      money   ['moni]

This pattern is particularly well-attested in varieties of Italian spoken in New York and Boston and other /r/-dropping areas of the United States where words ending in /r/ are pronounced with final schwa. They are adapted with a final /a/ and no gemination (although these words never undergo stress shift).

(14) *lover*   ['lova]
      shoe-maker   [ʃu'meka]
      teacher   ['titʃa]

There is another set of data containing a non-geminate consonant that could be analyzed in two ways. American English has a rule of flapping of intervocalic /t/ and /d/: *city* ['sɪɾi], and not *['sɪti] (as in, for example, British English). In the adaptation of English words containing a flap, we never find gemination.

(15) *city*   ['siɾi]
      water   ['vwoɾa]
      what's the matter? [vatstsa'maɾa]

The use of a singleton consonant in these cases might be due to the fact that English flap is most similar to the Italian singleton [ɾ] phoneme, along the lines of what we saw in (7) above. And a tap, by definition, is short. Alternatively, these data might pattern with the data in (13) and (14) above. An analysis of this latter approach is presented below in §5.4.

5.2   English v́cvc#

If the English word ends in a consonant, and stress is on the penultimate syllable, the Italian adaptation will have an additional final vowel, thereby adding a syllable, and stress will be on the antepenultimate syllable. Crucially, the consonant following the stressed vowel is not geminated.

(16) *shovel*   ['ʃabola]
      trouble   ['trobolo]
      people   ['pipoli]

This generalization is violated when the consonant following the stressed vowel is one of the 'inherently long' consonants discussed in §3 above.

(17) *bushel*   ['buʃʃolo]

5.3   English ́vc#

If the English word contains a final stressed VC syllable, and the final C is an obstruent (final sonorants will be discussed in §6 below), the adapted form contains an added final vowel, and the obstruent is geminated.

(18)   *bread*   ['breddi]
       beach   ['bitʃtʃa]
       mistake   [mis'tekka]
       roof   ['ruffo]
       book   ['bukko]

5.4   Analysis

All of the data in §5.1, §5.2, §5.3 can be accounted for in a unified way. First, we must posit the *Principle of Morphological Analysis of Borrowed Nouns* whereby a foreign noun is identified as an Italian stem (Repetti 2006).

(19)   foreign noun = native stem

The way in which the stem is incorporated into the phonological structure of Italian is determined by an alignment constraint active the loan adaptation process (Repetti 2006; Shinohara 2003). The constraint Align-R(stem, σ) requires the right edge of the stem (identified as the foreign noun) to be aligned with a syllable. This alignment constraint is part of the integration process, and is not part of the regular production grammar. Gemination, therefore, can be understood as a means of keeping the foreign stem separate from the Italian suffix, as illustrated in the data in §5.3. Cases in §5.1 in which the consonant is not geminated are due to the fact that the foreign stem is already aligned with a syllable, and gemination would be superfluous. The data in §5.2 are also immune to gemination despite the fact that they violate the alignment constraint. Gemination is blocked in these cases because of more highly ranked markedness constraints banning certain metrical structures, and in particular a heavy syllable following a stressed syllable.

(20)   Align-R(stem, σ) – the right edge of the stem (identified as the foreign noun)
       must be aligned with a syllable (Shinohara 2003; Repetti 2006)
       *GemCons — no geminate consonants

(21)   ranking: Align-R(stem, σ) >> *GemCons

These two constraints allow us to account for the data in §5.1–§5.3 in a straightforward way. (Remember that gemination in (17) is due to a fact about the inventory of Italian consonants: intervocalic /ʃ/ is always long, as discussed in §3.)

(22)

money	Align(stem, σ)	*GemCons
a. [ˈmo.ni] ←		
b. [ˈmon.ni]		*!

(23)

shovel	Align(stem, σ)	*GemCons
a. [ˈʃa.bo.la] ←	*	
b. [ˈʃab.bo.la]	*	*!

(24)

book	Align(stem, σ)	*GemCons
a. [ˈbu.ko]	*!	
b. [ˈbuk.ko] ←		*

The alignment constraint is not violated in (22) since the stem is vowel final and therefore aligned with a syllable. In this case, the markedness constraint eliminates the losing candidate.

In (23) both candidates violate the alignment constraint, and *GemCons again selects the winner. Another possible candidate, such as *[ˈʃab.ol.la], which does not incur a violation of Align(stem, σ), is eliminated by a higher ranked metrical constraint banning post-tonic heavy syllables.

In (24) the alignment constraint eliminates candidate (a) since the stem, /buk/, is not aligned with a syllable. Candidate (b), with gemination, does not violate this constraint and is, therefore, the winner, despite its violation of *GemCons.

6. Vowel tenseness in the English word

6.1   Data

There is one additional category of borrowings in which consonant length cannot be accounted for using the abovementioned principles. If the English word contains a final stressed VC syllable and the final consonant is a sonorant, the Italian adaptation contains an added final vowel, and the final sonorant may or may not be geminated. The choice between the geminate and singleton sonorant is determined by the tenseness of the preceding stressed vowel. I am assuming that the English tense vowels are [i u e o a ɑ], and the English lax vowels are [ɪ ʊ ɛ ɔ ʌ æ].

If the English vowel is lax, the sonorant is geminated in the Italian form. This is consistent with the patterns described in §5.3.

(25) *bill*   ['billo]
      ten   ['tenne]
      pull   ['pullo]
      son   ['sonni]

If the English vowel is tense, the sonorant is not geminated in the Italian adaptation.

(26) *green*   ['grini]
      bowl   ['bolo]
      high school ['ajskula]
      wholesale   ['olsele]
      lane   ['lena]

This pattern is regular with sonorant /l/ and /n/. We cannot test this pattern with words containing a final sonorant /r/ (in non-/r/-dropping varieties) because of historical changes in English resulting in the loss of vowel tenseness distinctions before /r/. The data from non-/r/-dropping varieties (such as Canadian Italian, Danesi 1985) all have a short sonorant.

(27) *Frigidaire*   [fridʒdʒi dera] (refrigerator)
      welfare   [wel'fera]
      hardware   [ard'weri]
      floor   ['floro]
      store   ['storo]

The presence of a short /r/ may be due to the realization of the vowel before /r/ as tense, in which case these data pattern with the data in (26) above. Alternatively, the lack of gemination may be due to a more general dispreference for high sonority geminates (see §6.2 and Footnote 3).

   The correlation between vowel tenseness and consonant length does not hold for data with sonorant /m/. It seems that the presence or absence of a geminate /m/ is unpredictable.

(28) a.   tense vowel + /m/ > [m]

            frame   ['frema]
            same   ['semi]
            steam   ['stima]
            ice cream   ['skrima]
      b.   tense vowel + /m/ > [mm]

            broom   ['brummi]
            game   ['gemma]
            team   ['timme]

c.   lax vowel + /m/ > [mm]

jam   ['dʒemma]
bum   ['bommo]

The data in (28a) with a tense vowel and a short sonorant, and the data in (28c) with a lax vowel and a long sonorant, behave as expected. But, how can we explain the data in (28b) with a tense vowel and a long sonorant? This may be due to the fact that in many central and southern varieties of Italian and Italian dialects, /m/ is always long in intervocalic position (Rohlfs 1966:§222). In other words, in some varieties, /m/ is an inherently long consonant similar to those described in §3. Hence, if there is a geminate [mm] in a borrowed form, it is not clear if it is due to the inherent length of the segment in a particular variety of Italian (28b)–(28c), or to the constraints in the integration process described in §5.4 above (28c). Unfortunately it is not possible to conduct a more detailed survey of these loans based on the regional origin of the Italian speakers since most of the data come from published sources which do not include this information. Very few native/fluent speakers of these American-Italian varieties exist, and most speakers are now heavily influenced by standard Italian.

     Our analysis does make a prediction about the distribution of singleton /m/. Since there are no varieties of Italian which permit only singleton /m/ (but allow other geminate consonants), if we find a singleton /m/ in intervocalic position, we should be able to explain its presence. Our analysis predicts that we should find singleton /m/ following a stressed tense vowel (although we might find a geminate here because of regional dialect influence), but we should not find singleton /m/ following a stressed lax vowels. This prediction is, in fact, borne out by the data.

(29)   lax vowel + /m/ > [m]
       unattested

Aside from the data involving bilabial nasals, the generalization seems to be that if you have a tense vowel in the English form, you cannot have a geminate sonorant in the Italian form. The chart below shows that consonant length is determined by vowel tenseness only with sonorants, not with obstruents.

(30)

	sonorant	obstruent
lax vowel	pull ['pullo]	foot ['futto]
tense vowel	high school [aj'skula]	suit ['sutto]

6.2   Analysis

How can we account for these patterns involving sonorants? Clearly, the Alignment analysis alone cannot work, nor can any of the other approaches outlined above.

In other languages we find similar patterns. For example, English tense vowels are adapted as long in loan words in Japanese, while English lax vowels are adapted as short (Katayama 1998; Abraham 2004).

(31) $V_{[-tense]} > V$
         $V_{[+tense]} > VV$

I propose that a similar process takes place in American-Italian. English vowel tenseness is mapped to Italian vowel length, which is included in the phonological representation that becomes the input form.[2] (See Abraham 2004; Jacobs & Gussenhoven 2000; Kabak 2003; Katayama 1998; Peperkamp & Dupoux 2003; etc. See also Nevins & Braun in this volume for another case of mapping an L2 output to an underlying representation which attempts to match the phonetics of the L2 form.)

A constraint forcing the quantity of the input vowel to be maintained in the output is necessary: Wt-Ident-I-O. This constraint must be ranked lower than the Align(stem, σ) constraint in order to allow for gemination of obstruents following tense vowels (see §5.3 and §5.4).

(32)   Wt-Ident-I-O – the weight of output vowels must be identical to their input correspondents (Kager 1999)

(33)   ranking: Align(stem, σ) >> Wt-Ident-I-O

Clearly, an additional factor is at play since vowel tenseness appears to be relevant in determining the length of the following sonorant but not the following obstruent.

Kawahara (2005) argues that geminate sonorants are cross-linguistically more marked than geminate obstruents, and that their markedness derives from the confusability of length contrasts for sonorants: "the more sonorous a segment is, the more difficult it is to perceive its segmental duration, and hence the less perceptible its geminacy contrasts are" (Kawahara 2005:1).[3] He proposes the

---

2.   A word here is in order regarding vowel length in Italian. Italian has a well-known process of vowel lengthening in stressed open syllables. Although the degree of lengthening may vary depending on the position of the stressed syllable in the word (penults lengthen the most), the generalization is that stressed vowels in open syllables are longer than stressed vowels in closed syllables (D'Imperio & Rosenthall 1999). These facts have been interpreted as a requirement on stressed syllables or head feet. I have not indicated vowel length in the previous data since it has not been relevant until this point.

3.   There are some areas of Italian grammar which can be interpreted as an attempt to avoid high sonority geminates and geminate [rr] in particular. In addition to the loan data illustrated in (27) in which only singleton [r] is attested, we find an avoidance of gemination of /r/ in "backwards raddoppiamento." ("Backwards raddoppiamento" is the lengthening of a

universal ranking in (34), which we can abbreviate as in (35) in which the more specific context banning geminate sonorant consonants is ranked more highly than the general context banning geminate consonants altogether.

(34) universal ranking:
　　　*GemGlides » *GemLiquids » *GemNasals » *GemObs

(35) *GemSon » *GemCons

The constraints requiring (i) the quantity of output vowels to be faithful to the input specifications (Wt-Ident-I-O), and (ii) the ban on geminate sonorants (*GemSon), are working together since geminate sonorants are banned only in an attempt to maintain the weight of the input long (tense) vowel. In other words, the markedness constraint (*GemSon) is activated only if the faithfulness constraint (Wt-Ident-I-O) is violated. In order to account for the data (26), we can posit a conjoined constraint, *GemSon & Wt-Ident-I-O, which eliminates a candidate output only if both of its conjuncts are violated (Lubowicz 2002; Prince & Smolensky 1993).

(36) *GemSon & Wt-Ident-I-O

This constraint must be ranked higher than Align(stem, σ) in order to block gemination in the data involving a tense vowel plus sonorant consonant, but not in the other forms.

(37) *GemSon & Wt-Ident-I-O >> Align(stem, σ)

We see how this constraint works in the following tableaux containing an input with a tense (long) or a lax (short) vowel followed by a sonorant or an obstruent.[4]

(38)

school /VV/	*GemSon & Wt-Ident-I-O	Align(stem, σ)	*GemCons
a. ['sku:.la] ←	√ & √	*	
b. ['skul.la]	* & * = *!		*

---

word-final consonant before a vowel-initial word.) Obstruents and non-/r/ sonorants lengthen in this context (*gas asfissiante* [gáss asfissiánte] 'asphyxiating gas,' *tram elettrico* [trámm eléttrico] 'electric tram'), but /r/ generally does not lengthen (*bar elegante* [bár elegánte] 'elegant bar'), although it may lengthen in polysyllabic oxytones (*bazar aperto* [badzdzárr apérto] 'open-air bazaar'). See Chierchia (1986) and Cardinaletti & Repetti (2008).

4.　I do not include candidates with a long vowel followed by a geminate consonant. Such candidates would be eliminated by the high-ranking *3μ constraint. I also avoided candidates with a short vowel followed by a singleton consonant since those candidates violate the metrical requirement on stressed syllables/feet (see Footnote 2).

(39)

pull /V/	*GemSon & Wt-Ident-I-O	Align(stem, σ)	*GemCons
a. [ˈpuː.lo]	√ & *	*!	
b. [ˈpul.lo] ←	* & √		*

(40)

suit /VV/	*GemSon & Wt-Ident-I-O	Align(stem, σ)	*GemCons
a. [ˈsuː.to]	√ & √	*!	
b. [ˈsut.to] ←	√ & *		*

(41)

foot /V/	*GemSon & Wt-Ident-I-O	Align(stem, σ)	*GemCons
a. [ˈfuː.to]	√ & *	*!	
b. [ˈfut.to] ←	√ & √		*

In (38a) neither of the conjuncts of the conjoined constraint is violated since the candidate contains a long vowel (like the input vowel), and it does not contain a geminate sonorant. Candidate (38b) violates both conjuncts of the conjoined constraint since it contains a short vowel (as opposed to the input) and a geminate sonorant. Since both conjuncts are violated, this candidate is eliminated, leaving candidate (38a) as the winner, despite the fact that it violates the lower ranked alignment constraint.

In Tableau (39) neither of the candidates violates both conjuncts of the conjoined constraint: (39a) violates the identity conjunct but not *GemSon, and (39b) violates *GemSon but not Wt-Ident-I-O. Therefore, the alignment constraint selects the winning candidate.

Finally, Tableaux (40) and (41) do not contain sonorants, so they cannot violate the *GemSon conjunct of the conjoined constraint. Hence, in these tableaux it is the alignment constraint that eliminates the loser.

7. Conclusions

In this paper, I hope to have shown how many aspects of grammar are involved in loan word adaptation, and specifically in the determination of consonant length.

(42)   –lexicon (similar native lexical items)
         –morphology (the identification of the 'stem')
         –phonetics-phonology (segment inventory, structure of stressed syllables, mapping of vowel tenseness to length, perceptibility of geminate sonorants)

If the foreign word is phonologically similar to a native word or if the foreign word ends in a series of segments that are similar to a native suffix, the length present in the Italian items is used in the adapted form. Morphological considerations such as the identification of the stem, and morpho-phonological considerations such as the alignment of the stem and a syllable, also affect the integration process. The phonological structure of the borrowing language, including the inventory of segments and constraints on syllable and metrical structure, interact in the determination of consonant length in loans. Finally, we saw that the mapping of vowel tenseness to vowel length and a constraint banning geminate sonorants are also involved in the integration of foreign loans.

References

Abraham, Diane. 2004. "Hearing Double: The Perception of Length Contrast in Loanwords". Ms., SUNY, Stony Brook.

Bafile, Laura. 2002. "Le consonanti finali nel fiorentino e nel napoletano". *Rivista italiana di dialettologia* 26.149–178.

Cardinaletti, Anna & Lori Repetti. 2008. "Phrase-Level and Word-Level Syllables: Resyllabification and the Prosodization of Clitics". *Phonological Domains: Universals and Derivations* ed. by Baris Kabak & Janet Grijzenhout. Berlin: Mouton de Gruyter, to appear.

Chierchia, Gennaro. 1986. "Length, Syllabification and the Phonological Cycle in Italian". *Journal of Italian Linguistics* 8.5–34.

Danesi, Marcel. 1985. *Loanwords and Phonological Methodology*. Ottawa: Didier.

D'Imperio, Mariapaola & Sam Rosenthall. 1999. "Phonetics and phonology of main stress in Italian". *Phonology* 16.1–28.

Di Vita, Rosario. 1931. "Corruzioni di vocaboli inglesi sicilianizzati". *Rivista italiana di letteratura dialettale* 3.266–273.

Haller, Hermann W. 1993. *Una lingua perduta e ritrovata: l'italiano degli italo-americani*. Firenze: La Nuova Italia.

Jacobs, Haike & Carlos Gussenhoven. 2000. "Loan Phonology: Perception, Salience, the Lexicon and OT". *Optimality Theory: Phonology, Syntax, and Acquisition* ed. by J. Dekkers, et al. 193–210. Oxford: OUP.

Kabak, Baris. 2003. *The Perceptual Processing of Second Language Consonant Clusters*. Ph.D. Dissertation, University of Delaware

Kager, René. 1999. *Optimality Theory*. Cambridge: CUP.

Katayama, Motoko. 1998. *Optimality Theory and Japanese Loanword Phonology*. Ph.D. dissertation, UCSC.

Kawahara, Shigeto. 2005. "Sonorancy and geminacy". ROA #755–0705.

Kenstowicz, Michael. 2003. "Salience and Similarity in Loanword Adaptation: A Case Study from Fijian". ROA #609–0803.

Livingston, Arthur. 1918. "La Merica Sanemagogna". *The Romanic Review* 9.206–226.

Lubowicz, Anna. 2002. "Derived Environment Effects in Optimality Theory". *Lingua* 112.243–280.

McCarthy, John & Alan Prince. 1993. "Prosodic Morphology I". Ms., University of Massachusetts, Amherst and Rutgers University.

Menarini, Alberto. 1939. "L'italo-americano degli Stati Uniti". *Lingua nostra* 1.152–160.

—— 1947. *Ai margini della lingua.* Firenze: Sansoni.

Mencken, H.L. 1947. *The American Language.* New York: Knopf.

Ortisi, Domenico. 1951. "Alcune osservazioni sulla formazione del dialetto siculo-americano". *Italica* 28.42–47.

Peperkamp, Sharon & Emmanuel Dupoux. 2003. "Reinterpreting Loanword Adaptations: The Role of Perception". Ms.,15th ICPhS, Barcelona.

Prince, Alan & Paul Smolensky. 1993. "Optimality Theory". Ms., Rutgers University and University of Colorado, Boulder.

Repetti, Lori. 1993. "The Integration of Foreign Loans in the Phonology of Italian". *Italica* 70.182–196.

—— 2003. "Come i sostantivi inglesi diventano italiani: la morfologia e la fonologia dei prestiti". *Italiano e inglese a confronto* ed. by Anna-Vera Sullam Calimani, 31–42. Firenze: Cesati.

—— 2006. "The Emergence of Marked Structures in the Integration of Loans in Italian". *Historical Romance Linguistics: Retrospective and Perspectives* ed. by Randall S. Gess & Deborah Arteaga, 209–239. Amsterdam & Philadelphia: John Benjamins.

Rohlfs, Gerhard. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti. Vol. I: Fonologia.* Torino: Einaudi.

Shinohara, Shigeko. 2003. "Emergence of Universal Grammar in Foreign Word Adaptations". Ms.

Tropea, Giovanni. 1978. "Sulla condizione dei dialetti italiani negli Stati Uniti". *La ricerca dialettale* 2.295–311.

Nasal harmony and the representation of nasality in Maxacalí

Evidence from Portuguese loans

W. Leo Wetzels

Université de Paris III, Sorbonne Nouvelle/LPP, CNRS/
Vrije Universiteit Amsterdam

Popovich (1971) claimed that nasality in Maxacalí is contrastive for both consonants and vowels. Rodrigues (1980) proposes to eliminate nasality entirely from the set of phonological features of Maxacalí and to represent all voiced, prenasalized and nasal consonants as voiced consonants in the lexical representation. In this paper, we take up this discussion and present an alternative analysis that is intermediate between the proposals by Popovich and Rodrigues. On the basis of evidence drawn from the adaptation of Brazilian loanwords in Maxacalí, it is concluded that nasality is contrastive for vowels.

1.   Introduction

Maxacalí, a member of the Macro-Jê linguistic family, is spoken in the border area of the states of Minas Gerais and Bahia, Brazil.[1] A 1997 census report from the *Instituto SocioAmbiental* suggests that there are roughly 802 Maxacalí. These people live in an area that is suffering from significant deforestation and over-grazing. The Maxacalí, who still teach the traditional language to their children, are largely monolingual, though some have a very rudimentary Portuguese ability. The survival of their language and culture is uncertain because of the precarious conditions under which they live; large-scale malnutrition arising from their belief that it is possible to survive by hunting and gathering without planting coupled with massive alcoholism call into question the long-term physical survival of this people. A detailed account of Maxacalí, and the most comprehensive

---

description to date as far as the phonology is concerned, became available with the publication of a paper by Gudschinsky, Popovich and Popovich in *Language* 46.1 (1970), henceforth referred to as GPP.[2,3] In their paper, GPP discuss the intriguing process of (pre)vocalization: in Maxacalí all coda consonants may develop a vocalic pre-articulation, which may take on the status of a full vowel and even entirely replace the consonant. In this study we will address the process of nasal harmony and the question of the representation of nasality in Maxacalí. We will compare two proposals that account for the distribution of nasal vowels and consonants in this language, one by GPP and one by Rodrigues (1980). Despite the radical differences in the underlying system posited in each of these proposals, both provide an observationally adequate account of the surface distribution of nasal sounds in Maxacalí. Thereafter, we will turn our attention to a set of data that appear to conflict with Rodrigues' analysis, because they indicate (at least a measure of) phonological contrast for vowels. In order to determine the relevance of these facts, we will look at how Maxacalí adapts loanwords from Portuguese that contain nasal vowels and/or nasal consonants. We conclude that the facts that conflict with Rodrigues' analysis do not represent a subset requiring exceptional lexical marking for vocalic nasality, but are indicative of a general nasal contrast for vowels.

---

2.   Drawing on material collected by Harold and Francis Popovich, Davis (1968) classified Maxacalí as a member of the Macro-Jê stock of South American indigenous languages. A more recent report on the classification of the Brazilian indigenous languages by Rodrigues (1986a) confirms Davies' hypothesis. Information about the history and the people of the Maxacalí can be found in Rubinger, Amorim & Marcato (1980). F. Popovich (1980) presents a study of the social organization of the Maxacalí.

3.   Almost all subsequent work focusing on the theoretical interpretation of the Maxacalí phonological facts is based directly or indirectly on GPP. McCawley (1967) and Hyman (1975:46, 191), make mention of the process of prevocalization. In Reighard (1972:540–1), glide formation and (pre)vocalization are briefly discussed. Drawing on Reighard (1972), Clements (1991) uses glide formation occurring between the underlying vowel and the (pre) vocalized coda consonant to support his proposal for defining place of articulation in consonants and vowels with a single set of articulator features. In Hume & Odden (1996), Maxacalí is briefly mentioned. A formal (feature- geometry) analysis of (pre)vocalization based on secondary consonant features was proposed by Wetzels (1993) and a slightly different analysis based on primary consonant features by Wetzels & Sluyters (1995). Araujo (2000) provides a more recent study of Maxacalí phonology and morphology (partly) based on independent fieldwork. Operstein's (to appear) cross-linguistic study of prevocalization also extensively discusses Maxacalí data, based on GPP, Wetzels (1993) and Sluyters and Wetzels (1995), reinterpreting Wetzels' (1993) analysis in a gestural framework.

2. Sources

Many of the examples on which our analysis is based are taken from Popovich & Popovich (2004). A further source of information is the unpublished paper Popovich (1983a). A more extensive discussion of Maxacalí prosody, with little attention for the segmental and syllable levels, is presented in Popovich (1985), a work that also contains 75 pages of phonetic transcription from which further examples were taken. Other examples were drawn from a study by Rodrigues (1980), which is based on data drawn from two unpublished papers by Popovich & Popovich (2004),[4] as well as from Popovich (1971) and GPP. Specifically with respect to the problem of nasality, we have also used an unpublished reaction to Rodrigues' analysis of Maxacalí nasality by Popovich (1983b) and an unpublished paper by Araujo (ms).

3. The consonants and vowels of Maxacalí

3.1 Vowels

Maxacalí distinguishes five oral vowels underlyingly, which are front high /i/, back high (non-labial) /ɯ/, front non-high /e/, back non-high (labial) /o/, and the low vowel /a/, all of which have a nasal correspondent. According to GPP, nasality is contrastive for vowels.

(1) Maxacalí Vowel Inventory

	front	central	back non-labial	labial
high	/i/		/ɯ/	
mid	/e/			/o/
low		/a/		

Both oral and nasal vowels show a wide range of surface variants.[5] The degree of aperture of the front high vowel /i/ varies freely from relatively close [i] to more open [ɪ] or even [e]: /tihic/ → [tɪhey] 'man'. Mid /e/ can be [e], but also more open [ɛ], or often even [æ]: /pepi/ → [pæpi] 'above'. The phonetic range of underlying /a/ extends from central low [a], to mid central [ə] and to very low

---

4. For which no references are provided.

5. In GPP (1970:80) the low vowel /a/ is assigned to the same aperture class as /e/ and /o/ at the contrastive phonological level. Here, we will treat /a/ as the sole member of the class of low vowels.

back [a]: /–kaʔok/ → [kaʔoᵘk̚] or [kaʔoᵘx] 'strong'; interestingly, the backmost variants of /a/ occur between alveolar consonants. The vowel /o/ is relatively stable, and invariably surfaces as back mid [o]. Lastly, back high /ɯ/ has the widest range of surface variants, ranging from high back to centralized mid front: /pipkɯp/ → [pipkɯˠp̚] or [pipkɪˠp̚] or [pipkɜˠp̚] 'nail'. In the processes of prevocalization and vocalization, other vowels may arise, as will be discussed in Section 3.2 below.

3.2   Prevowels

Maxacalí shows a large discrepancy between the lexical representation of words and the way in which they are actually pronounced. The most frequently cited phenomenon of Maxacalí sound structure is the seemingly free variation that exists between consonants and vowels at the end of syllables[6]:

(2)

    a.   /t/, /n/ are optionally pronounced as mid central [ɜ], [ɜ̃], respectively:
        /pat.kɯp/ ~ [paɜ.kiɤ] 'rib'
        /tõ.mãn/ ~ [tõ.mãɜ̃] 'tomato'

    b.   /p/, /m/ are optionally pronounced as the mid non-round back vowels [ɤ], [ɤ̃], respectively:
        /pap.tɯc/ ~ [paɤ.tɯi] 'drunk'
        /mi.hĩm/ ~ [mĩ.hĩɤ̃] 'wood'

    c.   /k/, /ŋ/ are optionally pronounced as the high, centralized back vowels [ɯ], [ɯ̃], respectively:
        /kɯcakkik/ ~ [kɯcaɯkiɯ] 'capybara'
        /ɲã.ɲã.mõŋ/ ~ [ɲã.ɲã.mõɯ̃] 'uncle has gone'

    d.   /c/, /ɲ/ are optionally pronounced as [i], [ĩ], respectively:
        /ko.kac/ ~ [ko.kai] 'lizard'
        /mã.ʔãɲ/ ~ [mã.ʔãĩ] 'alligator'

Maxacalí seems unique in allowing the complete set of syllable-final consonants to be realized as vowels. According to GPP, this (pre)vocalization is most prominent in 'stressed or prolonged syllables' (1970:82).[7] More concretely, in unstressed

---

6.   In the examples in (2) and below, it is provisionally assumed that nasality is phonological for both consonants and vowels. We will return to this issue later.

7.   Syllables are prolonged under emphatic stress as in calling or pedantic repetition (cf. GPP: 81). Also, as Harold Popovich points out (p.c.), the concept of 'augmentation' is signaled by stress and length realized either on the nuclei of all the syllables of the word or on the last syllable only.

syllables prevocalization never results in a fully-fledged vowel, whereas this does happen when the syllable is prolonged. If a coda consonant shares its active articulator (place of articulation) with the immediately following onset, be it word-internally or across word boundaries, it is always vocalic and very often no consonantal closure takes place in the coda. The quality of the vocalic element that develops is predictable from the nature of the underlying consonant, whether prevocalization or complete vocalization occurs. The surface qualities of the (pre)vowels derived from coda consonants are provided in (3):[8]

(3) Prevowel quality

	Front	Central	Back
High	i/_{c,ɲ}		ɯ/_{k,ŋ}
upper			ɤ/_{p,m}
Mid			
lower		ɜ/_{t,n}	

3.3 Consonants

Maxacalí has ten consonantal segments at the contrastive phonemic level. Observe the absence of liquids and (phonemic) fricatives. The system includes two laryngeals: the glottal stop /ʔ/ and the glottal fricative /h/. As for the supralaryngeal consonants, GPP posit two series of underlying segments, voiceless and nasal stops, both of which are realized at four distinctive places of articulation, as in (4) below (see GPP:80):

(4)

	Labial	Alveolar	Palatal	Velar	Laryngeal
Oral	p	t	c	k	ʔ, h
Nasal	m	n	ɲ	ŋ	

At the phonetic level, voiced non-nasal consonants and prenasalized consonants may occur, as in [kadop] from underlying /kanop/ 'to mix dry substance with liquid' and [ⁿɟokoba] from /ɲokoma/ 'below'. Both plain and prenasalized voiced stops are analyzed by GPP as surface variants of underlying nasals, which are denasalized partially or completely by contiguous oral vowels. GPP's analysis also posits a lexical contrast with respect to nasality for vowels, as in /pãhãm/ 'bread' versus /paptɯc/ 'drunk'.

---

8. All (pre)vowels derived from consonants are non-labial. When derived from a nasal consonant, the (pre)vowel always surfaces as nasalized. In both GPP and in Popovich (1983), the (pre)vowel appearing with velars is defined as back, whereas in Popovich (1985) it is considered central. In Wetzels & Sluyters (1995), it is classified as dorsal.

4.   Nasal harmony[9]

In Maxacalí, nasal sounds usually do not stand on their own, but co-occur with other nasal sounds, as in the words in (5):

(5)   [mãɲõn]   'sun'   [mõŋ]   'to go'
      [kõmãn]   'stepmother'   [nãŋ]   'fall (singular subject)'
      [ʔuĩŋãɲ]   'nervous'   [uĩmõŋ]   'choose'
      [muĩnuĩŋ]   'deer'   [ɲĩmãn]   'wing'
      [mãhãm]   'fish'   [pĩn(ẽŋ)]   'sound made by jumping'
      [pẽnãŋ]   'see'   [kõmẽn]   'city'
      [pãhãm]   'bread'   [ɲuĩmuĩŋ]   'all of us'
      [ɲãmĩɲ]   'spirit'   [mĩhĩm]   'wood'
      [muĩnuĩn]   'ant'   [ɲãnãm]   'light'

In the words above, the only non-nasal segments are voiceless consonants. Otherwise, all segments are nasal. Comparatively, the words in (6) are completely oral.

(6)   [tihic]   'man'   [potaʔ]   'to cry'
      [kaʔok]   'strong'   [pacok]   'corn'
      [kokec]   'dog'   [kot]   'manioc'
      [kabah]   'also'   [cokakkak]   'chicken'
      [codat]   'soldier'   [caʔ]   'to fall (plural subject)'
      [puɯtop]   'to bite'   [tapet]   'paper'
      [puɯtuɯc]   'heavy'   [capɯɯp]   'pig'
      [pepi]   'above'   [pohoc]   'arrow'
      [pihep]   'reservoir'   [(n)ɬokoba]   'below'
      [kadop]   'to mix in liquid' [topaʔ]   'to fly'

Maxacalí possesses a number of roots that contain disharmonic sequences, as in (7):

(7)   [kacuĩɲ]   'like this'   [puɯcõɲ]   'worm'
      [kiʔuĩn]   'parallel lines' [cuɯkãŋ]   'above'
      [patĩɲ]   'could be'   [kocãm]   'fish hook'
      [caʔãm]   'slug'   [tokãn]   'toucan'
      [kacõɲ]   'praise the lord' [kacĩɲ]   'thus'
      [muĩtik]   'with'   [ʔãᵐbuɯhuɯ]   'wind'

Under specific conditions, voiced consonants are optionally realized as prenasalized stops. This happens when the voiced stop appears word-initially as the onset of an oral syllable, as shown by the examples in (8) below:

---

9.   In this section and below, the effect of (pre)vocalization as well as other low level phonetic facts will be disregarded when irrelevant for the exposition.

(8) [⁽ᵐ⁾boj] (<Port *boi* [boj])   'ox'
     [⁽ⁿ⁾doʔok]   'wave'
     [⁽ⁿ⁾ɟokoba]   'below'
     [⁽ⁿ⁾gahap](<Port *garrafa* [gaˈhafa]) 'bottle'

The great majority of Maxacalí roots ends in a consonant and usually does not
contain more than two syllables. Nouns are mostly bisyllabic of the type CVCVC,
whereas function words as well as verbs may be monosyllabic. Words longer than
two syllables can be derived by compounding, a very productive process in Maxacalí.
Polysyllabic words may also be derived by affixation, mostly suffixation. Prefixes
are rare. The majority of the suffixes start with a voiceless consonant; very few start
with a voiced or nasal consonant or with a vowel. Affixes may be inherently nasal
or oral. In derived words, nasal roots or affixes freely combine with oral roots, as
is illustrated in (9):[10]

(9) [puɨtuɨc+nãŋ]   bird+diminutive   'little bird'
     [mĩm+tuɨt]   wood+structure   'house'
     [capuɨp+nãŋ]   pig+diminutive   'white-lipped peccary'
     [hãm+cop+bac]   work+collective+good   'good things'
     [mãɲõŋ+hec]   sun+female   'moon'
     [mĩn+cop]   leaf+collective   'leaves'
     [mĩm#koc]   wood#hole   'canoe'
     [tik#kuɨm]   man#sister   'daughter'
     [mĩm#koc#ɟok]   wood#hole#straight   'straight canoe'
     [ĩm+pata#cac]   my+foot#skin   'my toe nail'
     [ĩm+pa#ce]   my+eye#hair   'my eyelashes'

So far, we have seen that in Maxacalí voiced [b,d,ɟ,g], voiceless [p,t,c,k], and nasal
[m,n,ɲ,ŋ] consonants occur on the phonetic surface. Word-initially as well as
word-internally, the voiced and voiceless consonants contrast in the onset of syl-
lables that contain an oral nucleus, whereas the nasal consonants may not occur in
the onset of such syllables:

(10) p   [ta*p*et]   'paper'   b   [ka*b*ah]   'also'
      t   [puɨ*t*uɨc] 'heavy'   d   [ka*d*op]   'scatter'
      c   [*c*itet]   'spotted' ɟ   [*ɟ*okoba]   'down'
      k   [*k*apec]   'coffee'   g   [*g*oc]   'to ambush'

---

10.   Through this work, the symbol + represents an affix boundary, # a compound-internal
boundary, and ## a word boundary.

On the other hand, in the onset of syllables containing a nasal nucleus, voice-less stops contrast with nasal consonants, whereas non-nasal voiced stops are not allowed in this position:

(11)   p   [*p*ãm]   'bread'   m   [*m*ã?ãɲ]   'alligator'
       t   [*t*õmãn]   'tomato'   n   [*n*ãmtuɾt] 'bow'
       c   [ka*c*ũɲ]   'like this'   ɲ   [*m*ãɲõn]   'sun'
       k   [*k*ãɲã]   'snake'   ŋ   [*ŋ*õɲ]   'to smoke'

The surface distribution of voiceless, voiced and nasal onset consonants in the native vocabulary of Maxacalí is summarized in (12) below:

(12)   Co-occurrence restrictions between nucleus and onset

$$
\begin{array}{cc}
\sigma & \sigma \\
/\,| & /\,| \\
\{p,b\}\ V & \{p,m\}\ V \\
| & | \\
[\text{oral}] & [\text{nasal}]
\end{array}
$$

The distribution of consonants in the syllable coda is even more restricted: after an oral nucleus only voiceless consonants occur, while after a nasal nucleus only nasal consonants are permitted, as exemplified in (13) below:[11]

(13)   p   [tu*p*]   'new'   m   [ãmõ*m*]   'Amom (name of a spirit)'
       t   [ãbo*t*]   'sand'   n   [ũ(hũ)*n*]   'Maxacalí woman'
       c   [caki*c*]   'to die'   ɲ   [cũ*ɲ*]   'to suffer'
       k   [ãpa*k*]   'to hear about'   ŋ   [ɲĩmã*ŋ*]   'wing'

The distribution of coda consonants is provided in (14):

(14)   Co-occurrence restrictions between nucleus and coda

$$
\begin{array}{cc}
\sigma & \sigma \\
|\,\backslash & |\,\backslash \\
V\ p & V\ m \\
| & | \\
[\text{oral}] & [\text{nasal}]
\end{array}
$$

---

11.   This is true at some intermediate level of representation, which we consider to represent the output of the lexical phonology. Phonetically, partial or complete denasalization of the coda consonant may occur in the sequence C]$_{\sigma\sigma}$[C-voice, when the coda and onset conso-nants do not share their place of articulation: compare [mĩ(hĩ)m] 'wood' and [mĩnta] 'fruit' with [mĩptut] 'house' and [mĩpkoj] 'canoe'. Equally, voiceless consonants are optionally voiced before nasal consonants, as in /taknõɲ/ → [tagnõɲ] 'brother'.

Since voiceless stops contrast with both voiced and nasal stops in onsets, voiceless stops must be distinguished as a separate phonological class. In contrast, in the syllable onset, voiced and nasal consonants are in complementary distribution, suggesting that one class may be derived from the other. As we have seen, non-nasal voiced stops do not occur in the syllable coda. In this position, voiceless and nasal stops are in complementary distribution.

As mentioned previously, GPP assume that in Maxacalí nasality is distinctive for both consonants and vowels. Voiced oral stops that occur in the onset of syllables with an oral nucleus are derived by a rule of consonant denasalization: /penec/ → [pedec] 'tree frog'. Partial denasalization derives nasal-oral contour segments from nasal consonants in word-initial onsets when followed by an oral vowel: /nac/ → [^{n}day] or [day]. In the onset of a syllable containing a nasal vowel and in codas after a nasal nucleus, nasal consonants remain nasal.

4.1 GPP versus Rodrigues

We note that in all the words given in (5), as well as in the individual morphemes contained in the derived words in (9), a nasal span ends at its right edge in a nasal consonant that is located in the syllable coda. It is this observation that was at the basis of Rodrigues' (1981) analysis of Maxacalí nasalization, to which we turn next.

In a reaction to GPP and Popovich (1971), Rodrigues (1980) made the interesting claim that nasality in Maxacalí is predictable for both consonants and vowels given that nasal spans almost invariably end in a word-final nasal consonant and voiced non-nasal stops and nasal consonants are in complementary distribution. He proposes the elimination of nasality from the set of Maxacalí contrastive features and the representation of all voiced, prenasalized and nasal consonants as voiced consonants in the lexical representation. According to Rodrigues, voiced stops are obligatorily nasalized when they occur at the right edge of the word, and it is from this position that they spread their nasal feature from right to left within the word domain, until blocked by a voiceless consonant. He also allows for a limited use of contrastive nasalization in vowels in cases where the nasal quality of the vowel cannot be obtained from a following nasalized consonant, as in /hã/ 'manner' vs. /ha/ 'is'. Prenasalized consonants are considered positional, word-initial variants of underlying voiced consonants.

Compared to the position taken by GPP, Rodrigues' proposal leads to a drastic simplification of the lexical representations of Maxacalí words, such that non-sonorant voiced and voiceless consonants freely occur in both onsets and codas of syllables, which contain only oral vowels. His analysis explains why nasal syllables cannot contain voiced non-nasal obstruents, as the latter would either trigger nasality projection (coda) or, along with the vocalic nucleus, be a target for nasal spreading (onset). Furthermore, the fact that all underlying vowels are

oral explains why only voiceless stops can surface as oral in the coda of syllables containing an oral nucleus. Finally, it is predicted that voiced and voiceless consonants may contrast in the onset before oral vowels, but that voiceless stops must contrast with nasal consonants before a nasal nucleus. In other words, all of the positional variation in the surface consonant system represented in (12) and (14) fall out naturally from the analysis.[12] Thus, it is possible to summarize Rodrigues' proposal as follows:[13]

(15)   *A voiced consonant in the syllable coda is obligatorily realized as a nasal sonorant which spreads its nasality leftward within the word-domain, until it is blocked by an intervening voiceless consonant* (/p,t,c,k/).[14]

In Maxacalí, the notions of 'word' and 'morpheme' overlap to a large extent. Morphemes, especially nouns, are most commonly bisyllabic and CVCVC-type. Words that consist of a single closed syllable $C_xV_iC_y$[15] regularly have a corresponding bisyllabic form $C_xV_i?V_iC_y$ or $C_xV_ihV_iC_y$; the short variant occurs when it is part[16] of a larger morphological structure and the long form is used when the word appears in isolation. Compare, for example, the noun phrase in (16a) with the compound in (16b):[17]

---

12.   The different positions adopted by GPP and Rodrigues (1981) are not only due to a difference in theoretical paradigm, i.e. Pikean phonemics vs. classical generative phonology, but are also related to a difference in objective. Whereas GPP, and especially Popovich (1983), are interested in developing a writing system for Maxacalí, Rodrigues (1980) approaches the problem from a strictly phonological perspective, aiming his description at a redundancy-free lexical representation of Maxacalí words.

13.   Compared to the original proposal by Rodrigues, the spreading principle stated in (15) differs with regard to the structural position of the trigger of nasality. According to Rodrigues (1980), the (voiced) consonant triggering nasalization must be word-final. Even if the notion 'word' is understood in the sense defined below, this would imply that non-derived words of the type (C)V̄NTVC, i.e. a hypothetical form *[mãɲtup], with a morpheme-internal nasal coda consonant, should not exist. Although such words are usually composite, we wish to leave open the possibility that not all words of this type are synchronically transparent. We have therefore preferred to rephrase the spreading principle in such a way that it refers to the end of syllables rather than to the end of words. For our specific purpose, which is to test Rodrigues' hypothesis concerning the underlying source for surface nasality in Maxacalí, this modification is not crucial.

14.   The sound [ʔ] likely also serves to block nasal spread. The behaviour of the glottal consonants is elaborated in Section 4.2 below.

15.   Subscripts indicate segment identity or the lack thereof.

16.   In the examples we have encountered, the short form is always the left element in the sequence. We suppose that this is a further condition on its distribution.

17.   The examples in this paragraph are adapted from Araujo (2000).

(16) a. $_\text{NP}$[mīhīm##koj]   wood##hole   'perforated wood'
      b. $_\text{N}$[mīm#koj]   wood#hole   'canoe'

Compare also the sequences in (17a–b), where the suffix /+tɛ/ represents the ergative marker:

(17) a. [tihik##mõɲõn]   the man##sleep   'the man sleeps'
      b. [tik+tɛʔ##cānãhãŋ]   the man+ERGATIVE##call 'the man calls'

In Maxacalí, the great majority of bisyllabic words in which the first syllable is closed are either compound structures, as in (16b); structures containing suffixes, such as [tik+tɛʔ] in (17b); or prefixed words, like /iN+pata/ 'my+foot' → [īm+pataʔ].

As is shown by the sequences /ce+nãŋ/ → [cēnãŋ] 'hair+DIMINUTIVE', or /īN+pa+nãŋ/ → [īm+pã+nãŋ] 'my+eye+DIMINUTIVE', nasality that originates in a suffix spreads leftwards onto the preceding base. This means that in Maxacalí the domain for nasal spreading is not the morpheme, but a larger structure. The question then arises as to whether nasality that originates on the right part of a compound propagates into the part(s) on the left. Vowel-final root morphemes are relatively rare in Maxacalí. Nevertheless, some compounds of the relevant type were encountered, which seem to show that the nasal feature does not spread across a compound boundary, as in /kuɲnãmãm/ [kuɲãmãm] 'light, lamp', which very probably contains the root /ku(hu)/ 'firewood'. While it could be argued that, in this word, nasal spreading is blocked by an underlying voiceless stop – /kuhu/ has a variant /kuhuk/ in Maxacali' – the presence of a blocking consonant is more difficult to justify for the verb [penãhã] 'to look', where [pe] most likely represents the noun /pa/ 'eye'. We will therefore assume that the notion 'word' in the statement (15) refers to a bare or affixed lexical category.[18]

The operation of the principle in (15) is illustrated with the following examples (periods indicate syllable boundaries):

(18) a. /ba.ɟod/ → [mã. ɲõn] 'sun'. In this word, syllable-final /d/ is realized as [n] and spreads its nasality leftward. Since there is no voiceless consonant in the word, nasality spreads over the whole sequence.
      b. /pɯ.coɟ/ → [pɯ.cõŋ] 'worm'. Here, syllable-final /ɟ/ is realized as [ŋ]. Nasality spreads leftward until it is blocked by voiceless [c].
      c. /tot.cic.pec/ → [tot.cic.pec] 'watermelon'. This word does not contain a syllable-final voiced consonant. Consequently, it surfaces with oral segments only.

---

18. From the phonetic transcriptions in Popovich (1985), we conclude that nasal spreading across compound and word boundaries may happen in informal speech styles. Unfortunately, a detailed account of Maxacalí morphology and a fine-grained analysis of the stylistic factors involved in nasal harmony are still lacking. We will therefore limit the discussion to obligatory nasal harmony, which, in the theoretical frame used here, corresponds with the lexical part of the harmony process (see also Footnote 11).

When considered exclusively from a phonological perspective, Rodrigues' analysis looks superior to the one proposed by GPP. It almost entirely eliminates nasality from the lexical representation. Both the range of surface distributions of voiced, voiceless and nasal stops, as well as the nasal/oral variation in vowels results from the free distribution of voiced and voiceless stops in lexical entries and from the projection of nasality in the syllable coda combined with the leftward propagation of the nasal feature.

   GPP's proposal, on he other hand, must posit a wholesale underlying oral/nasal contrast for consonants and vowels. It must account for optional prenasalization and obligatory consonant denasalization and also include some spreading mechanism to explain alternations between oral and nasal vowels in words that contain oral vowels underlyingly, such as /ce/ 'hair', which surfaces with a nasal vowel when combined with the diminutive suffix /nãŋ/:[cēnãŋ]. Both the elegance of the proposal and its capacity to account for the great majority of Maxacalí words have led us to adopt Rodrigues' analysis in our earlier studies of Maxacalí phonology,[19] despite the fact that, as far as we are aware, Maxacalí would be the only language known in which surface nasality is derived for both consonants and vowels from underlying representations in which nasality is entirely lacking. The existence of a seemingly small residue of words for which nasality could not be derived from a coda consonant, such as /hã/ 'manner' vs. /ha/ 'is', did not seem alarming because the generalization expressed in (15) holds for all words that have a nasal, i.e. an underlyingly voiced coda consonant. Moreover, one could conjecture that the original conditioning factor for the nasality in words like /hã/ had become lost in the current state of the language, but was present at some earlier stage.

   In an unpublished reaction to Rodrigues' analysis, Popovich (ms.) pointed out that the set of words containing nasal vowels that cannot be derived by the principle in (15) is relatively large.[20] The example sets (19a–b), taken from Popovich (*ibid*) and other sources contain such words:

(19) a. [hãmãⁿdiccok] 'play'   [ʔãᵐbɯc]   'needle'
         [hãᵐbedac]   'sell'   [ʔãᵐbɯɯh]   'wind'
         [kãɲãⁿdok]   'viper species' [ʔãᵐbɯk]   'to cook'
         [kãmãⁿdok]   'horse'   [ʔãⁿɟihik]   'non-Indian'

     b. [hõmãh]   'remote time'   [ɲõkãmũŋ]   'side'
         [nãnã]   'uncle'   [nõã]   'finish'
         [mãʔac]   'will eat 3p SG' [mãkak]   'heron'

---

c. [mãɲõn]   'sun'   [mõŋ]   'to go'
    [kõmãn]   'stepmother'   [nãŋ]   'fall (singular subject)'
    [ʔuĩŋãɲ]   'nervous'   [ũmõŋ]   'choose'
    [muĩnuĩŋ]   'deer'   [nõʔõm]   'that one'
    [mãhãm]   'fish'   [pĩn]   'sound made by jumping'
    [pẽnãŋ]   'see'   [kõmẽn]   'city center'
    [pãm]   'bread'   [mãʔãɲ]   'alligator'
    [ɲãmĩɲ]   'spirit'   [mĩ(hĩ)m]   'wood'

d. [kiʔũn]   'parallel lines'
    [caʔãm]   'slug'
    [kɯʔĩn]   'bands'
    [mãʔãɲ]   'alligator'

The words in (19a) are different from the ones given in (5), repeated here as (19c), to the extent that in the examples of (19c) all coda consonants can be (pre)vocalized, whereas the nasal consonants that are transcribed with superscript symbols in the words in (19a) never are. Since in Maxacalí all coda consonants may develop a (pre)vowel, (pre)vocalization provides us with a reliable test for determining the syllable position of a segment. The fact that the relevant nasal consonants in (19a) cannot develop a (pre)vowel shows that they are not located in the coda, but represent the nasal part of a single prenasalized onset segment (created by progressive spreading of nasality from the preceding nasal vowel). The words in (19a–b) demonstrate that nasal vowels may occur inside a word without the presence of a nasal coda from which their nasality can be derived.

    The words in (19d) deserve special attention, because the glottal stop [ʔ] exhibits contradictory behavior as it blocks nasal spreading in the first two examples but not in the last one. As observed earlier (cf. 16/17), monosyllabic $[C_xV_iC_y]$-type morphemes regularly have a corresponding bisyllabic form $[C_xV_iʔV_iC_y]$ or $[C_xV_ihV_iC_y]$, where the short variant occurs as a part of a larger morphological structure, while the long form is used when the word appears in isolation. Rodrigues (1986:82) suggests that the long forms are derived from the short ones, proposing, for example, /bib/ as the underlying form of [mĩhĩm]. In Wetzels & Sluyters (1995), following GPP and Popovich (1985:45), it was proposed that short forms were derived from long forms, because the occurrence of the type of glottal sound, [ʔ] or [h], was believed to be unpredictable. However, according to Araujo (2000), the glottal sound that occurs in the long forms, although it is usually realized as [h], alternates freely with [ʔ]. Some examples are given in (20):

(20) [mãhãm]   ~ [mãm]   'fish'
      [mĩhĩm]   ~ [mĩm]   'wood'
      [mãhãɲ]   ~ [mãɲ]   'alligator'
      [pohok]   ~ [pok]   'marsh'
      [tuhuɯt]   ~ [tuɯt]   'bag'

In (21) the way in which the segments of the short forms correspond with those of the long forms is shown for the morpheme [mãɲ] ~ [mãhãɲ] 'alligator':

(21)   {h,ʔ}

                       |

         C V C V C

          \ \ / /

             m ã ɲ

             | | |

          C   V   C
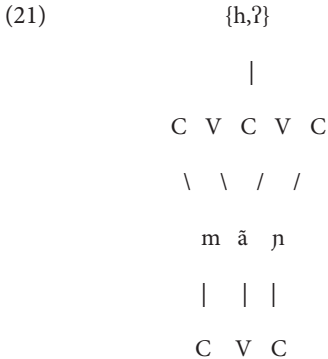
Araujo's observation that the glottal stop and the fricative are in free variation in these words is important, since it allows us to circumvent the question as to whether the long form is derived from the short form or vice-versa. Given the predictability of the glottal sound in the long forms, we can explain the 'transparency' of the glottal sounds irrespective of the directionality of the derivational process. If the short forms are the lexical base from which the long forms are derived, the words in which these consonants look transparent result from the lengthening process illustrated in (21), not from spreading: /CVC/→/CV{h,ʔ}VC/, with the mapping of the melody to the unspecified C and V positions. If the long form is the starting point of a shortening process, we may assume a mechanism like /CV$_x$V$_x$C/→[CV$_x$C], and a rule of default onset insertion between V$_x$V$_x$ in the long form. In both scenarios, we do not need to assume that glottal sounds are transparent to nasal harmony.[21]

4.2 The nativization of BP loans in Maxacalí

Considering the examples in (19a–b), one is led to conclude that, for almost all the nasal vowels in these words, nasality cannot be derived from a nasal coda consonant. The question thus arises as to whether these words suggest a different

---

21. The fact that not all monosyllabic words have long forms could be interpreted as an argument in favour of a shortening process. However, there are also words of the type XV$_i${h/ʔ}V$_i$C that have no short forms. Interestingly, BP loans which satisfy the conditions for shortening are never shortened, whereas monosyllabic BP morphemes are usually lengthened, a fact which suggests the existence of a lengthening rule: BP *garrafa* [gaˈhafa] 'bottle' > Max /gahap/ [ᵑgahaˠp], never *[ᵑgaˠp]. Compare this with BP [ˈʃav(i)] *chave* 'key' > Max /cahap/ [tʃahaˠp]. As will be illustrated in Section 4.2, the latter word is the lengthened version (/cahap/) of an intermediate */cap/, after deletion of the word-final vowel in the BP input sequence. In contemporary Maxacalí the only form in use is /cahap/, which cannot be shortened.

analysis of nasality in general; for example, one in which coda consonants are not triggers but targets of nasal spreading. In order to shed light on this question, we will turn to a set of Maxacalí words that are of Brazilian Portuguese (BP) origin. In particular, we will see how BP words that contain nasal vowels and/or nasal consonants are integrated in the Maxacalí sound system. Accordingly, a short excursus into the phonology of nasality of BP is necessary.

The BP system of consonantal phonemes is much more complex than the Maxacalí one, as table (22) shows:

(22) BP Consonantal Phonemes

Labial	Alveolar	Palatal	Velar
p	t		k
b	d		g
f	s	ʃ	
v	z	ʒ	
m	n	ɲ	
l		ʎ	
	ɾ		x

The lateral [l] is pronounced [ɯ] in the syllable coda of the large majority of BP dialects. The alveolar tap [ɾ] and the velar fricative [x] are in complementary distribution, except intervocalically[22] where they contrast: *caro* 'expensive' ['kaɾu]↔ ['kaxu][23] *carro* 'car'. Word-initially, in the syllable coda, and syllable-initially following a consonant, only [x] occurs: ['xatu] *rato* 'rat', [max] *mar* 'sea', ['ʒeNxu] *genro* 'son-in-law'. The tap occurs as the second element of a complex onset, a position from which [x] is banned: ['kɾizi] *crise* 'crisis'. In the BP dialect that is spoken in the area where the Maxacalí live, [x] is pronounced [h]. The coronal stops /t, d/ are realized as the affricates [tʃ,dʒ] before [i]. In BP, nasal consonants contrast with both voiced and voiceless non-nasal stops in the onset of oral and (surface) nasal syllables: /paR/[24] *par* 'pair', /baR/ *bar* 'bar', /maR/ *mar* 'sea'; /paNda/ ['pãdɐ] *panda* 'panda', /baNda/ ['bãdɐ] *banda* 'band', /maNda/ ['mãdɐ] *manda* 'sends-3p'.[25] In the syllable coda, only sonorant sounds (glides, liquids, N) and /s/ are allowed.

---

22. In Wetzels (1997), the contrast between /ɾ/ and /x/ is represented phonologically as one between /R/ and /RR/, respectively, unspecified for place features. In positions where there is no contrast, /R/ is posited.

23. In these examples and henceforth, the symbol ' marks the following syllable as having primary word stress.

24. /R/ is predictably realized as either [x] or [ɾ], depending on its syllable position: *mar* 'sea' [max] ~ [maɾɪs] *mares* 'sea-PL'.

25. In the preceding examples, N represents a nasal mora not specified for place of articulation (see Wetzels (1997) for extensive discussion of nasality in BP).

BP has seven contrastive vowels under primary stress /i, e, ɛ, a, ɔ, o, u/. In unstressed position or when nasalized, the distinction between upper and lower mid vowels is neutralized. The phonetic quality of unstressed oral mid vowels is subject to dialectal variation, ranging between upper and lower mid, whereas nasal mid vowels are usually pronounced as upper mid. Two types of nasality must be distinguished for vowels. The traditional division is between 'allophonic nasality' and 'contrastive nasality', although, for both types, nasality is predictable from a following nasal consonant. Allophonic nasalization targets vowels before syllable-initial nasal consonants, as in /banana/ 'banana', which is pronounced [baˈnɐ̃nɐ] or [bɐ̃ˈnɐ̃nɐ], depending on the dialect.[26] Contrastive nasality results from the obligatory spreading of the [nasal]-feature from a nasal consonant (or mora) in the syllable coda to the preceding nuclear vowel, as in /kaNpo/ [ˈkɐ̃pɷ] *campo* 'countryside' or /paNkada/ [pɐ̃ˈkadɐ] *pancada* 'hit'.[27] Moreover, BP has a number of nasal diphthongs, the most frequent of which is [ɐ̃ɷ̃], as in /televizaũ/ [televiˈzɐ̃ɷ̃] *televisão* 'television'. Nasality thus appears to be a pervasive feature of both the source and the borrowing language. Nevertheless, the differences in the surface distribution of both oral and nasal sounds are considerable. As we have just seen, BP shows a free distribution of the different consonant types (oral or nasal) in the onset. Also, the nasality of a syllable nucleus is not contingent on a (surface) nasal coda, as in the first syllable of [ˈkɐ̃.pɷ] < /kaNpo/ *campo* 'countryside'. On the assumption that the productive constraints of Maxacalí phonology are somehow visible in the process of nativization of BP words, the way in which the sounds and sound sequences of BP are adapted to the phonotactic structure of Maxacalí should allow us to verify the correctness of the rules proposed by GPP and Rodrigues.

Above it was observed that the Maxacalí consonant system lacks liquids and supralaryngeal fricatives. When these sounds occur in BP loans, liquids and /f, v/ are changed into plosives, whereas BP [s, ʃ, tʃ, z, ʒ, dʒ] are usually pronounced as affricates, such that [s, ʃ, tʃ] correspond to Maxacalí [tʃ] and BP [z, ʒ, dʒ] correspond

---

to Maxacali [dʒ] in contexts where these affricates exist as (prevocalic) allophones of the palatal stops /c, ɟ/. In the syllable coda, these segments are borrowed as stops. In the process of nativization, BP consonants that are foreign to Maxacalí preserve their place of articulation. BP [ɾ, l] become [t, d, n] and [f, v] become [p, b, m], while the choice of laryngeal and manner features is to a large extent determined by their position in the Maxacalí syllable and the nature of the syllable nucleus (be it oral or nasal). Interestingly, the nucleus from a BP loan in Maxacalí will be nasal if the BP source has a nasal nucleus and/or a nasal onset. As it turns out, the laryngeal (voiced or voiceless) specification of the BP coda consonant never accounts for the oral or nasal character of the syllable nucleus in the corresponding Maxacalí word. The words in (23) illustrate the different ways in which BP [f,v] are adapted in Maxacalí. The starred examples involving the variables X and/or Y represent sequences that contain [f, v] in contexts not attested in the borrowed vocabulary. They show how we would expect these consonants to be adapted based on the treatment of other BP consonants, as demonstrated below.

(23)　a.　BP [f]　　　　　　　　　　～　Max [p]

　　　　i.　before oral and nasal nucleus

　　　　　　fogão　　　[fuˈgʒɶ̃]　　　～　[puˈⁿgãm]　　'stove'
　　　　　　*XfṼY　　　　　　　　　～　*XpṼY

　　　　ii.　after oral nucleus

　　　　　　garrafa　　[gahaf(ɐ)]　　～　[ⁿgaˈhap]　　'bottle'

　　　b.　BP [f]　　　　　　　　　　～　Max [m]

　　　　　　after nasal nucleus

　　　　　　*XṼf　　　　　　　　　　～　*XṼm

　　　c.　BP [v]　　　　　　　　　　～　Max [⁽ᵐ⁾b]

　　　　　　before oral nucleus

　　　　　　vaqueiro　[vaˈkejɾ(ɷ)]　　～　[⁽ᵐ⁾baˈket]　　'cowboy'
　　　　　　canivete　[kaniˈvɛtʃ(i)]　～　[kɯdiˈbɛt]　　'pocketknife'

　　　d.　BP [v]　　　　　　　　　　～　Max [m]

　　　　i.　before nasal nucleus

　　　　　　*XvṼY　　　　　　　　　～　*XmṼY

　　　　ii.　after nasal nucleus

　　　　　　*XṼv　　　　　　　　　　～　*XṼm

　　　e.　BP [v]　　　　　　　　　　～　Max [p]

　　　　　　after oral nucleus

　　　　　　chave　　　[ˈʃav(i)]　　　～　[tʃaˈhap][28]　　'key'

28.　The word [tʃaˈhap] is a lengthened form. See Footnote 21.

In the process of nativization, the BP word-final vowel is usually deleted when unstressed, which is why these vowels are put in parentheses in the BP examples above and below. When the BP word ends in a nasal diphthong, the final vowel is often interpreted as the vocalized surface manifestation of an underlying nasal consonant, as can be observed for the words *fogão* 'stove' [fuˈgɜ̃ɷ̃] → Max [puˈⁿgãm] or [puˈⁿgã̃ʳm], as well as in *calção* 'shorts' [kaɷˈsɜ̃ɷ̃]→ Max [kotˈtʃãm]. Also, a word-final syllable that is stressed in BP may be closed with a consonant in Maxacalí, as in *café* 'coffee' [kaˈfɛ] → Max [kaˈpec]. The corresponding pairs in (23) show that the labial fricatives of BP are nativized as labial stops. The laryngeal specification is preserved in the process of borrowing except when it conflicts with the phonotactic constraints of Maxacalí. This is clear from the example in (23e), where, in the absence of any language-specific distributional restrictions, one would expect *[tʃaˈhab] instead of [tʃaˈhap]. In conformity with the constraint expressed as (14) above, only voiceless stops may close a syllable containing an oral nucleus, which explains why [b] is borrowed as its voiceless counterpart [p] in this word. This example is particularly revealing, because, on the basis of Rodrigues' analysis of nasality, one would expect the voice character of /b/ in the hypothetical form */cab/ to be crucial in the adaptation process, as one observes for word-initial voiced consonants, as in [ᵐboj] < BP *boi* 'ox' or [ⁿgahap] < BP *garrafa* 'bottle', [ⁿdʒeˈdʒɯjʃ] < BP [ʒeˈzujs] *Jezus*, where voicing automatically triggers prenasalization. We would therefore expect BP *chave* [ˈʃavi] to appear as *[cãm] (or *[cãhãm] by lengthening), which would be in full agreement with the phonotactic requirements of Maxacalí, but, from the perspective of the oral/nasal distinction, in the opposite manner. As it turns out, the determining factor in the nativization of the word [ˈʃav(i)] is the orality of the syllable nucleus, not the voicedness of the coda consonant. The same treatment of BP voiced consonants can be observed in the words provided in (24), where he relevant segments appear in boldface:

(24)

				BP		Maxacalí
a.	Maxacalí	'Maxacalí'		[maʃakaˈli]	~	[mātʃakaˈdij]
	Gabriel	'Gabriel'		[gabɾiˈɛɷ]	~	[gabidiˈet]
	relógio	'watch'		[heˈlɔʒ(iɷ)]	~	[heˈdoc]
	retrato	'picture'		[hetɾat(ɷ)]	~	[hetaˈdat]
b.	martelo	'hammer'		[mahˈtɛl(ɷ)]	~	[ᵐbahˈtɛt]
	vaqueiro	'cowboy'		[vaˈke(j)ɾ(ɷ)]²⁹	~	[ᵐbaˈket]
	(na)feira	'market'		[naˈfe(j)ɾ(ɐ)]	~	[nãˈpet]

---

29.   The optional monophthongization of the diphthongs /ej/ and /aj/ is a process that occurs throughout Brazil. It happens most frequently before /ɾ/, but also before the palatal fricatives

c.	compadre	'godfather of my child'	[kũˈpad**r**(i)]	~	[koˈpa**t**]
	soldado	'soldier'	[soɷˈda**d**(ɷ)]	~	[tʃoˈda**t**]
	Margarida	'Margarida'	[mahgaˈ**r**id(ɐ)]	~	[mãᵑgaˈdi**t**]
	capado	'castrated pig'	[kaˈpa**d**(ɷ)]	~	[kaˈpa**t**]

The examples in (24a) show that [l, ɾ] appear as voiced coronal stops in the onset of syllables with an oral nucleus. When these sounds occur in the coda after an oral vowel they are devoiced, in conformity with constraint (14), as can be seen in (24b). As expected, the same occurs with syllable-final BP [d], as in the words in (24c).

We conclude that the nativization of BP words in Maxacalí shows that nasality in this language is not derivable from the [+voice] specification of a (non-nasal) coda consonant. We must therefore consider other options, for example, that nasality is a lexical property of consonants, as proposed by GPP. In this analysis, the underlying representation of a word like [mũĩnuĩɲnãmtit] 'goat' would be /mɯnɯɲnamtit/. Nasal coda consonants trigger nasal harmony and nasal onsets are transparent for spreading. Since leftward spreading is independently necessary to account for alternations like [ce] ~ [cẽnãŋ], this possibility is worth considering. However, under this analysis, it would still not be possible to account for the nasal vowels in words that have no coda, such as *nãnã* 'uncle', *nõã* 'finish', and others, as in (19a,b) above. One may next consider the hypothesis that nasality spreads bidirectionally: right-to-left, to account for words like *pãm* 'bread' and for alternations like [ce] ~ [cẽ+nãŋ]; and left-to-right, from a nasal onset to the following nucleus (and coda) in words like *nãnã* 'uncle', *nõã* 'finish'. Left-to-right spreading is moreover suggested by the borrowings in (25) below, where we observe that when a BP syllable has a nasal onset, even in the presence of an oral nucleus, the syllable containing it appears as entirely nasal in Maxacalí, while eligible BP voiced and voiceless consonants indiscriminately appear as nasal coda consonants.

(25)

		BP		Maxacalí
comércio	'city center'	[kõˈmɛhs(iɷ)]	~	[kõˈmɛ̃**n**]
tomate	'tomato'	[tũˈmatʃ(i)]	~	[tõˈmã**n**]
janela	'window'	[ʒĩˈnɛl(ɐ)]	~	[tʲĩˈnẽ**n**]
caneta	'ballpoint'	[kãˈnet(ɐ)]	~	[kãˈnẽ**ɲ**]
carneiro	'sheep'	[kahˈne(j)ɾ(ɷ)]	~	[kahˈnẽ**n**]
remédio	'medicine'	[heˈmɛd³(iɷ)]	~	[hehˈmẽ**n**]

Yet, one would not wish to conclude on the basis of the way the BP words in (25) are adapted that nasality is a lexical property of onset consonants in Maxacali.

/ʃ, ʒ/, and more frequently in stressed syllables than in unstressed ones (for discussion and references, see Wetzels 2007).

The reasons for rejecting this hypothesis are several. One is that there would be no explanation for the nasal vowels in words like those below:

(26) [hãᵐbedac] 'sell'
      [ʔãᵐbuɾk] 'wind'

In the words in (26) there is no nasal onset (nor a nasal coda[30]) from which nasality could spread. Furthermore, if nasal onset consonants were underlying and were allowed to spread their nasality to their right, it would not be possible to account for the complementary distribution between voiced non-sonorant consonants, which occur exclusively in the onset of oral syllables, and nasal consonants, which are prohibited in this environment (cf. words like *kabah* 'also', *kadop* 'scatter', and the words in (26), among many others). As a matter of fact, the voiced consonants in these words, if derived from an underlying nasal consonant, show that there can be no rightward nasal spreading, because the spreading of nasality to the tautosyllabic vowel would bleed the rule of consonant denasalization. Let us therefore consider the possibility that nasality is contrastive in vowels.

The BP words that are part of the corresponding sets provided below contain at least one nasal vowel or diphthong.

(27)   BP   Maxacalí
    a. quinhentos 'five hundred' [kī'ɲēt(ɸs)] ~ [kī'ɲēn]
    b. macarrão   'pasta'   [maka'hȝɶ̃]   ~ [mākā'hām]
                                                      (= /mã'+kām/?)
    c. feijão   'beans'   [fej'ȝȝɶ̃]   ~ [pē'ɲõŋ]
    d. tucano   'toucan'   [tu'kān(ɸ)]   ~ [to'kān]
    e. calção   'shorts'   [kaɸ's ȝɶ̃]   ~ [kot't ͡ʃ ãm]
    f. santo   'saint'   ['sāt(ɸ)]   ~ [t ͡ʃ ãn]
    g. compadre   'godfather of my child' [kũ'ᵐpadɾ(i)] ~ [kõm'pat]

From the point of view of the distribution of nasality, the BP word in (27a) represents a perfect Maxacalí word, except for the voiceless consonant following the nasal nucleus, which must be changed into the corresponding nasal stop. The BP word in (27b) is adapted to Maxacalí in a way that suggests that [h] is transparent. However, this conclusion might not be warranted, because the word could be interpreted as being derived from a monosyllabic morpheme underlyingly, as indicated. The activity of nasal spreading is particularly observable in (27c), where the nasal feature that is restricted to the word-final vowel in the BP noun is spread throughout the corresponding noun in Maxacalí, affecting all segments except the initial voiceless consonant. The word pairs in (27d–g) show the blocking effect of

30. For the prenasalization of [b] in these examples, recall our earlier discussion following the example set in (19a).

voiceless stops. We would be able to account for all of the words in (27), on the assumption that nasality is a lexical feature of vowels alone, by appealing to two mechanisms. The first assures that the nasal feature of the syllable nucleus spreads to the syllable coda, and the second mandates that the nasal feature of the nucleus spreads leftward to consonants of a specific type, as well as to vowels. These mechanisms are illustrated below for the word *pẽɲõŋ* 'beans'[31]:

(28)

$$
\begin{array}{ccccc}
\text{p} & \text{e} & \text{ɟ} & \text{o} & \text{K} \\
| & | & | & | & | \\
\text{C} & \text{V} & \text{C} & \text{V} & \text{C}
\end{array}
$$

= [pẽɲõŋ]

rhyme

[nasal]
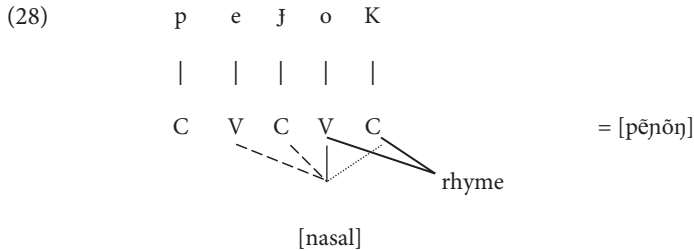
Next, we return to the words in (25), repeated as (29), and another set of loans, not discussed so far, provided in (30):

(29)

		BP		Maxacalí
comércio	'shop'	[kõˈmɛhs(iɰ)]	~	[kõˈmɛ̃n]
tomate	'tomato'	[tũˈmatʃ(i)]	~	[tõˈmãn]
remédio	'medicine'	[rẽˈmɛdʒ(iɰ)]	~	[hɛ̃ˈmɛ̃n]
janela	'window'	[ʒĩˈnɛl(ɐ)]	~	[tʃĩˈnẽn]
caneta	'ballpoint'	[kãˈnet(ɐ)]	~	[kãˈnɛ̃ŋ]
carneiro	'sheep'	[kahˈne(j)ɾ(ɰ)]	~	[kahˈnɛ̃n]

(30)

		BP		Maxacalí
Maxacalí	'Maxacalí'	[maʃakaˈli]	~	[mãtʃakaˈdij]
Marisa	'Marisa'	[maˈɾizɐ]	~	[mãⁱⁿdiʒa]
mesa	'table'	[ˈmezɐ]	~	[ˈmẽⁿdʒa]
moto	'motorbike'	[ˈmɔtɰ]	~	[mõˈtok]
moça	'girl'	[ˈmosɐ]	~	[ˈmõtʃaʔ]
(no) posto	'post'	[noˈpost(ɰ)]	~	[nõˈpoc]
(na) feira	'market'	[naˈfejɾ(ɐ)]	~	[nãˈpet]

In the words of (29), nasal consonants are word-internal, whereas in (30), they are word-initial. Together, the words in (29–30) show that the presence of a nasal onset in the BP words is sufficient to yield a fully nasal syllable in the corresponding

---

31. The capitals ɟ and K in (28) represent partially specified consonants. The question of the lexical specification will be taken up below.

Maxacalí words.[32] This comes as a surprise if Maxacalí has a productive rule that denasalizes nasal consonants in the onset of oral syllables, as was suggested by GPP. To be clear, there are some rare examples in which BP nasal onsets are oralized in Maxacalí, as in the examples below:

(31)   BP   Maxacalí
        martelo   'hammer'   [**mah**ˈtɛlɐ]   ~   [**⁽ᵐ⁾bah**ˈtɛt]
        canivete   'pocketknife'   [ka**ni**ˈvɛtʃi] ~   [ku**di**ˈbɛt]

The fact that similar BP words are nativized in different ways shows that the borrowing process is primarily oriented towards making BP surface patterns unknown to Maxacali compatible with the phonotactics of that language. The words in (29–31) do not necessarily show that either denasalization or left-to-right nasal spreading from syllable onsets are productive processes of Maxacalí phonology, but only that these are alternative strategies to repair ungrammatical surface sequences. Clearly, the preferred strategy is the nasalization of the oral vowel that is tautosyllabic with the BP nasal onset. In other words, confronted with a BP syllable containing a nasal onset and an oral nucleus, the Maxacalí speaker interprets the nasal onset as an indication of the nasality of its nucleus. This is what one would expect if nasality is contrastive in vowels and if spreading occurs from right to left. Note also that in the words in (30), only the vowel immediately to the right of the nasal consonant appears as nasal in Maxacalí. It is clearly not the case that the nasal feature is 'set afloat' in order to dock on the rightmost nasalizable segment in a sequence that does not contain a voiceless consonant. Were this true, we would expect, for example, the Maxacalí form corresponding to BP [maˈɾizɐ] 'Marisa' to be [mãˈnĩɲã], rather than [mãⁱⁿdiʒa]. The attested outcome for [maˈɾizɐ] is the one we would expect in view of the fact that in Maxacali a nasal span may be followed by an oral span, as in the words /muːtik/ 'with', [ʔãᵐbuɦuɪ] 'wind', etc. The repair that is applied is minimal, but sufficient to bring the sequence [maˈɾizɐ], and others like it, in agreement with the phonotactic requirements of the language.

5.   A different analysis of nasality in Maxacalí

From the way in which BP words containing nasal segments are adapted to the sound structure of Maxacalí, we have concluded that nasality is a contrastive feature

---

32.   Notice that given the words in (30), the allophonic nasality of the pretonic vowels in the BP words in (29) becomes irrelevant for the explanation of the nasality in the corresponding Maxacalí words. It could be claimed that in the Maxacalí words the complete nasal span is derived from the (hypothesized) nasality of the stressed vowel, by coda nasalization and leftward spreading.

of Maxacali vowels. We will show next that this hypothesis is sufficient to account for the surface distribution of this feature in both vowels and consonants in the native and borrowed vocabulary of Maxacalí.

In the analysis proposed below, we assume that predictable features are not part of the underlying phonological structure, but are provided when necessary in the course of the phonological derivation or by the phonetic component (cf. Clements, 2001). We will moreover dispense with unmarked features and feature specifications as long as they are not phonologically active in the grammar or in a specific part of the grammar that is, for independent reasons, distinguished from other parts.

We have seen above that Maxacalí has a series of voiceless consonants and a series of phonemes that has both voiced non-nasal and nasal consonants as its surface allophones. Taking into account the fact that nasal consonants surface only as onsets and codas of nasal nuclei, nasality for consonants is a predictable feature. Voiced and voiceless consonants contrast as onsets of oral syllables and vowels contrast for nasality. Furthermore, Maxacalí has no liquids or fricatives. Disregarding place of articulation in consonants and vowels as well as aperture distinctions in vowels, the remaining contrastive features of Maxacalí are presented as in (32):

(32)   Consonants   Vowels

              [−vocoid]   [+vocoid]
              ⌒   ⌒
    [−voice]   [+voice]   [−nasal]   [+nasal]
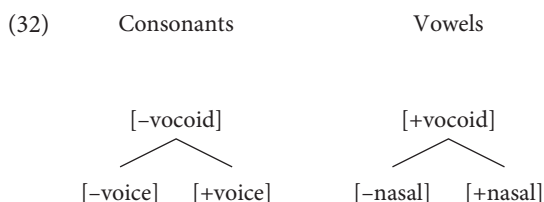
Notice that the major category features [approximant] and [sonorant] are not distinctive in Maxacalí. Moreover, in the syllable coda the opposition between voiced and voiceless consonants is neutralized: a [−vocoid] segment in the coda of a syllable containing an oral nucleus will be realized as a voiceless stop, whereas in the coda of a syllable containing a nasal nucleus, it will surface as a nasal consonant. Dispensing with the unmarked and phonologically inactive features [−voice] and [−nasal], we can set up a lexical representation in which syllables that surface with a nasal vowel are lexically represented as either (33a) or (33b), and syllables that surface with an oral vowel as either (34a) or (34b), where C represents a [−vocoid] segment, V a [+vocoid] segment:

(33)   a.   C   V   C   b.   C   V   C
                | | |
            [+voice] [+nasal]   [+nasal]

                  [mãn]   [pãn]

(34) a. C V C   b. C V C

[+voice]

[bat]   [pat]
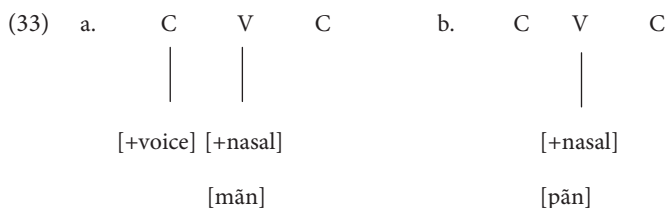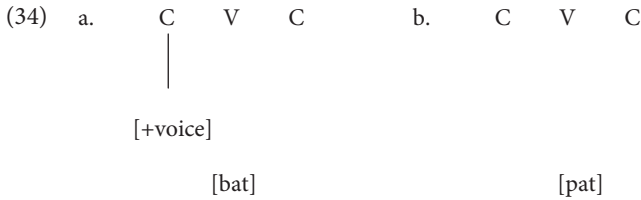
To derive the proper distribution of surface nasal consonants and vowels, we must posit two spreading mechanisms. One, indicated with a dotted association line in (35), spreads nasality from left to right and assures that syllable rhymes are always entirely oral or entirely nasal; and the other, indicated with a striped association line in (35a), iteratively spreads nasality from right to left predicting that possible targets to the left of a nasal vowel always surface as nasal.

(35)   rhyme   rhyme

      a.   C   V   C   b. C   V   C

[+voice] [+nasal]   [+nasal]

[mān]   [pān]
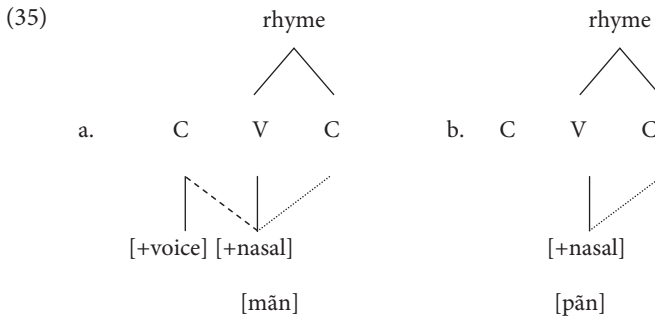
The proposed analysis explains why words with a nasal syllable immediately preceded by an oral vowel, nasal syllables with a voiced non-nasal onset, and syllable rhymes that disagree in nasality do not occur. However, an explicit analysis of Maxacalí nasality must account for more. If nasal vowels really had a free distribution, we would expect to find, apart from words like #kacuɨɲ# 'like this', #muɨtik# 'with', or #mãɲõn# 'sun', words of the type #ṼC$_1$ṼC$_2$#, where C$_1$ is a voiceless consonant. The non-occurrence of such (non-derived) words shows that each morpheme contains only a single instance of the [+nasal] feature. It follows that morphemes containing more than a single nasal sound are the result of [+nasal]-spreading from a unique segmental source, a nasal vowel in the case of Maxacalí.[33]

---

33. Some surface exceptions to this generalization exist, as the word [mõcãn] 'to arrive (plural subject)'. However, these words can generally be shown to consist of more than one morpheme. For the case at hand, compare [mõcaha] 'to arrive', [mõcakuc] 'to enter (plural subject)', [mõkpok] (</mõŋpok/) 'to send', [mõⁿgãhã] 'to lead'. From these and other words we establish the presence of a morpheme /mõŋ/, which also occurs independently in Maxacalí, meaning 'to go'.

In light of the preceding discussion, we establish the relevant parameters that account for the surface distribution of nasality in Maxacalí as given in (36):

(36)   Relevant parameters for Maxacalí nasal harmony[34] [35]

trigger	nasal vowels	nasal vowels
domain	word, as defined earlier	syllable rhyme
direction of spreading	left-to-right	(right-to-left)
target condition	[C, +voice][34]	C
spreading mode	iterative	(non-iterative)
adjacency	strictly local *N / \\ VCV	(strictly local)
OCP	* N N (domain: 'morpheme')	
feature co-occurrence restrictions[35]	*ŋ	none

In (36) we follow Peng (2000) and Clements & Osu (2001) for the set of relevant spreading parameters, except that we have defined a target condition for consonants, whereas these authors have chosen for the definition of a class of blocking segments. Nasal harmony in Maxacalí has all the properties of a lexical phenomenon. Our choice to define a class of targets instead of blockers is a consequence of our decision not to specify phonologically inactive features and because, in this part of the phonology of Maxacalí, no rule appears to refer to the feature [−voice]. The absence of vowels from the definition of the set of target segments follows from the general observation that when a language activates a spreading constraint for the nasal feature and when voiced consonants are targets, all segments that are higher on the sonority scale are included in the class of target segments. The conjunction of the locality condition which prohibits non-adjacent nasal consonants, with the OCP, which restricts [+nasal] to one instances per morpheme explains the ungrammaticality of hypothetical (morphologically underived) *#kãcuɪ̃ɲ# on

---

34.   Notice that, in an analysis that would specify voiceless segments with a lexical [−voice] feature, it is possible to limit the target of nasal spreading to just C and to obtain the blocking effect of voiceless consonants through a feature co-occurrence restriction *[+nasal, −voice].

35.   Interestingly, [ŋ] does not usually occur as the onset of a nasal nucleus, instead [ⁿg] is found: [ⁿgõɲ] 'to smoke'.

the assumption that well-formed nasal spans are dominated by a single [+nasal] feature. Since the nasality in both syllables of *#kãcũ̄ɲ# cannot result from spreading - the intervocalic voiceless consonant blocks nasal harmony - this word could only be derived from two tauto-morphemic instances of the nasal feature, a situation prohibited by the OCP. Notice, finally, that the parameters concerning the spreading of nasality in the syllable rhyme are either irrelevant or derivable from the domain specification (rhyme).

6. Discussion and conclusion

Maxacali is one of the numerous indigenous South American languages that lack an opposition between non-sonorant voiced stops and plain nasal consonants. Instead, this language establishes a phonetic relation between voicing and nasality through a process of nasality spreading from contrastively nasal vowels to voiced segments.[36] In this paper, we have shown that the hypothesis of a phonological contrast between nasal and oral vowels is necessary and sufficient to account for the distribution of nasality at the phonetic surface. The plausibility of the proposed analysis was reinforced by the manner in which BP words are borrowed into Maxacali, in particular by the way in which BP consonants were made to fit the severe restrictions on the shape of Maxacali coda consonants, by reference to the orality or nasality of the nuclear vowel.

(37)　…ṼN##
　　　…ṼT##　　…**VT**##
　　　…ṼD##　　…**VD**##

Of the most frequent input sequences to the Maxacalí nativization procedure given in (37),[37] only the boldfaced ones have a direct correspondent in Maxacalí phonotactics. The other three must be adapted for two reasons: First, because nasality vs. orality is homogeneous within the (surface) rhyme; and second, because voiced consonants only occur as onsets of oral syllables. If nasality were predictable from the contrastive voice feature in coda consonants in a system without phonological nasal vowels, we would expect BP voiced consonants in this position to be realized

---

36. The same surface co-occurrence of voice and nasality is observed in the optional process of word-initial prenasalization of voiced stops followed by an oral vowel.

37. Recall that in the BP input these word-final sequences usually end in a vowel which is deleted in the borrowing process. A BP sequence …VN(V)## could only be exceptionally part of the input to the nativization process because stress is predominantly prefinal in BP, and, consequently, the left vowel would be obligatorily nasal.

as nasal. One could object to this view by referring to the fact that adaptation strategies do not always mimic the rule system of the language, and that the change VD → VT could be an alternative (and maybe even simpler) way of repairing the illicit input structure. If this were true, one would have to explain why ṼT## and ṼD## are not nativized by undoing the nasality of the vowel, a feature which is non-contrastive in a grammar where nasality is derived from non-sonorant voiced consonants. On the other hand, the hypothesis of contrastive nasality for vowels allows us to straightforwardly predict how all BP rhymes are adapted to the nasal/oral categorization of Maxacalí on the basis of the existing underlying vocalic contrasts and rules of the language. We therefore conclude that the nasal segments and sequences in native words that cannot be derived from voiced coda consonants are not exceptions but counterexamples to the claim that surface nasality in Maxacalí is derived from non-sonorant voiced consonants in the coda.

We have found no cases of words borrowed from BP in which coda consonants are adapted to fit the distributional restrictions of Maxacalí without reference to the orality/nasality of the tautosyllabic vowel.[38] This is not always the case for BP nasal onset consonants as there are examples that remain unadapted. We also observed that sometimes a nasal onset consonant is oralized instead of the nucleus being nasalized, though this occurs exclusively in unaccented syllables. The BP words found in our loanword corpus in which nasal consonants are onsets of unstressed syllables appear in (38) below.

(38)

		BP		Maxacalí
martelo	'hammer'	[**ma**hˈtɛlɷ]	~	[⁽ᵐ⁾**ba**hˈtɛt]
canivete	'pocketknife'	[**ka**niˈvɛtʃi]	~	[kɯ**di**ˈbɛt]
Maxacalí	'Maxacalí'	[**ma**ʃakaˈli]	~	[**mã**tʃakaˈdij]
Marisa	'Marisa'	[**ma**ˈɾizɐ]	~	[**mã**ⁿdiʒa]
Margarida	'Margarida'	[**ma**hgaˈɾid(ɐ)]	~	[**mã**ⁿgaˈdit]
macarrão	'pasta'	[**ma**kaˈhɜ̃ɷ̃]	~	[**mã**kãˈhãm] (= /mãˈkãm/)
(na)feira	'market'	[**na**ˈfe(j)ɾ(ɐ)]	~	[**nã**ˈpet]
(no) posto	'post'	[**no**ˈpost(ɷ)]	~	[**nõ**ˈpoc]

In most unstressed syllables, Maxacalí follows the same nativization strategy as in stressed syllables. Few examples were encountered in which a nasal onset was oralized, i.e., BP /NV/ becomes Maxacalí /DV/ instead of /NṼ/. It is not entirely clear why the first two words in (38) follow a different adaptation pattern. Maybe the slight nasalization of the postnasal vowel in BP /NV/ syllables is actually fully

---

38. Directly, as in *santo* [ˈsãt(ɷ)] → [tʃãn] 'saint', or indirectly, as in *tomate* [tũˈmatʃ(i)]→ [tõˈmãn] 'tomato', through the reanalyzed sequence /toˈbãC/ with a nasal vowel.

perceived by the native speakers of Maxacalí[39] in stressed syllables, but less so in unstressed syllables. Another possibility for the recurrent nasalization of the stressed vowels as opposed to the less frequent nasalization of unstressed vowels could be the perceptibly higher saliency of the nasal onset in stressed syllables. In unstressed syllables either of the strategies could be equally likely, the preference for the nasalization of the nuclear vowel being an artifact of our small sample.

Independently of what the correct explanation is, the way *martelo* 'hammer' and *canivete* 'pocketknife' are nativized shows that Maxacalí speakers are aware of the vocalic oral/nasal contrast of BP words, even in contexts in which it does not occur in their own language. Similarly, the non-sonorant voiced consonants are perceived as such by Maxacalí speakers before nasal vowels, where they do not occur in Maxacalí words, as in the words listed in (39):

(39)		BP		Maxacalí	
feijão	'beans'	[feˈʒɐ̃ʊ̃]	~	[peˈʒõŋ]/[pẽ̄ɲõŋ]	
televisão	'television'	[televiˈzɐ̃ʊ̃]	~	[tedebiˈdˀãm]	*[tẽnẽmĩˈɲãm]
fogão	'stove'	[fuˈgɐ̃ʊ̃]	~	[puⁿgãm]	
laranja	'orange'	[laˈɾɜ̃ʒɐ]	~	[ⁿdaˈdãɟ]	*[nãˈnãɟ]
Lourenço	'Lourenço'	[loˈɾẽsɐ]	~	[ⁿdoˈdĩn]	*[nõnĩn]
João	'João'	[ʒɐ̃ʒɐ̃ʊ̃]	~	[ʒõˈãm]	*[ɲõˈãm]

In borrowings like BP *sabão* [saˈbɐ̃ʊ̃] → Max [tˠãˈmãm] 'soap', BP *feijão* [feˈʒɐ̃ʊ̃] → Maxacalí [pẽɲõŋ] 'beans', or BP *comércio* [kõˈmɛhs(iɐ)] → Maxacalí [kõˈmɛn] 'shop', sequences of voiced segments undergo R→L nasal harmony just as in the native vocabulary of Maxacalí. Aside from the expected [pẽɲõŋ] for 'beans', the form [peˈʒõŋ] also exists and a number of words were encountered that, given leftward harmony, are nativized irregularly. Indeed, the words in (39) show that non-sonorant voiced stops create exceptions to nasal harmony in BP loans.

From the discussion above, we conclude that nasal consonants, voiced oral consonants, and oral vowels are interpreted as such by Maxacalí speakers even in contexts in which they do not occur in Maxacalí. From this perspective, it is relevant to recall that exceptions to rightward spreading do not exist. This discrepancy may be due to the fact that in Maxacalí segments like /p, b/ contrast in the onset of oral syllables, while those like /p, m/ contrast in the onset of nasal syllables (at least superficially). On the other hand, consonants in the syllable coda only contrast for place of articulation. In other words, in the syllable onset, speakers are trained

---

39. João Moraes (p.c.) pointed out that this slight nasalization of postnasal vowels really exists in BP without being perceived by BP native speakers. It is unclear if the degree of phonetic nasality is different in stressed and unstressed syllables.

to recognize the cues distinguishing /p, b, m/ as relevant to the interpretation of their corresponding lexical phonological categories, which is not the case for the manner and laryngeal features that occur in the syllable coda, where voicelessness and nasality are purely articulatory categories. Should this interpretation prove correct, this would provide evidence for the non-specification analysis proposed above.

BP sounds are directly mapped onto the very limited set of Maxacalí distinctive features and thus distinctions like [±approximant], [±sonorant] and [±continuant] that are non-contrastive in Maxacalí are ignored. During the nativization process, specific features may be deleted or transferred to force illicit sequences to conform to Maxacalí phonotactics. In the BP word *martelo* [**ma**hˈtɛlɐ] → [⁽ᵐ⁾**ba**hˈtɛt] ʻhammerʼ, BP [ma] is (exceptionally) reanalyzed as underlying /ba/ in Maxacalí, whereas, generally, BP [NV] is reanalyzed lexically as Maxacalí /DṼ/. In the syllable coda, BP consonantal distinctions other than place of articulation are ignored and realized according to the requirements of the phonological grammar of Maxacalí. As to the nativization of non-sonorant voiced onsets, exceptions are created suggesting that in the borrowed vocabulary the set of target segments for leftward nasal harmony has become restricted to vowels. To the extent that the number of such ʻirregularʼ loans increases and these words come to be perceived as genuine Maxacalí words , the contact with BP will introduce a contrast between /p, b, m/ into the language.

References

Araujo, Gabriel. 2000. *Fonologia e Morfologia da Língua Maxakalí*. MA thesis. Universidade Estadual de Campinas (UNICAMP), Brazil

Araujo, Gabriel. 2005. *A Nativização de Emprestimos in Maxacalí*. ms.

Clements, George, N. 1991. Place of Articulation in Consonants and Vowels: A Unified Theory. *Phonetic and Phonological Studies on Vowel Features. Working Papers of the Cornell Phonetics Laboratory*, no. 5:77–121.

Clements, George N. & Sylvester Osu. 2001. *Nasal Harmony in Ikwere: Contribution to a Typology of Nasalization.* Paper read at the 3rd International Conference of the GDR (phonologie), Nantes, 2001. Handout.

Clements, George, N. 2001. Representational Economy in Constraint-based Phonology. T. Alan Hall (ed.) *Distinctive Feature Theory*. Berlin: Mouton de Gruyter.

Davis, Irvine. 1968. Some Macro-Jê Relationships. *International Journal of American Linguistics*, vol. 34:42–47.

Gudschinsky, Sarah, Harold Popovich, & Frances Popovich. 1970. Native Reaction and Phonetic Similarity in Maxacalí Phonology. *Language*, 46, 77–88.

Hume, Elizabeth & David Odden. 1996. Reconsidering consonantal. *Phonology* 13, 3:345–376.

Hyman, Larry, M. 1975. *Phonology, Theory and Analysis.* New York, Holt, Rinehart & Winston.

McCawley, James. 1967. Le rôle d'un système de traits phonologiques dans une théorie du language. *Langages*, 8:112–123.

Operstein, Natalie. 2008. *A new model of intrasegmental consonant structure: Evidence from consonant prevocalization and related processes.* Ms. University of Pittsburgh.

Peng, Long. 2000. Nasal Harmony in Three South-American Languages. IJAL 66.1:76–97.

Popovich, A. Harold. 1971. *The Sun and the Moon, a Maxacalí Text.* Estudos sobre Línguas e Culturas Indígenas. SIL. Brasília.

Popovich, A. Harold. 1983a. *Como um Som Maxacalí Pode Influenciar os Sons Contíguos.* Apostila para o Seminário de Cultura Indígena na Universidade Federal de Juiz de Fora, 28 de abril de 1983. Unp. ms.

Popovich, A. Harold. 1983b. *Resposta à 'Nasalização e Fronteira de Palavra em Maxacalí'.* Unp. Ms.

Popovich, A. Harold. 1985. *Discourse Phonology of Maxacalí. A Multilevel, Multiunit Approach.* Unp. MA Thesis, University of Texas at Austin.

Popovich, A. Harold & Frances Popovich. 2004. *Dicionário Maxacalí-Português com Glossário Português-Maxacalí.* SIL: Cuiabá, MT, Brazil.

Popovich, Frances. 1980. *The Social Organization of the Maxacalí.* MA Thesis. Sociology Department, The University of Texas at Arlington. Portuguese version available via SIL, Brazil.

Reighard, John. 1972. Labiality and Velarity in Consonants and Vowels. *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*: 533–543.

Rodrigues, Aryon Dall'Igna. 1980. *Nasalização e Fronteira de Palavra em Maxacalí.* Paper read at the V Encontro Nacional de Lingüística, PUC, Rio de Janeiro.

Rodrigues, Aryon Dall'Igna. 1986a. *Línguas Brasileiras. Para o Conhecimento das Línguas Indígenas.* São Paolo, Edições Loyola.

Rodrigues, Aryon Dall'Igna. 1986b. Silêncio, Pausa e Nasalização. Anais do VIII Encontro Nacional de Lingüística. Rio de Janeiro.

Rubinger, Marcos Magelhães, Maria Stella de Amorim, & Sonia de Almeida Marcato. 1980. *Índios Maxacalí: Resistência ou Morte.* Belo Horizonte, Interlivros de Minas Gerais Ltda.

Wetzels, W. Leo. 1993. Prevowels in Maxacalí: Where they come from. *Revista da Associação Brasileira de Lingüística*, São Paulo: 39–63.

Wetzels, W. Leo. 1995. Oclusivas Intrusivas em Maxacalí. Leo Wetzels (ed.). *Estudos Fonológicos de Língüas Indigenas Brasileiras.* Editora da UFRJ. Rio de Janeiro: 85–102.

Wetzels, W. Leo. 1997. The Lexical Representation of Nasality in Brazilian Portuguese. *Probus, 9.2*:203–232.

Wetzels, W. Leo. 2007. Primary Stress in Brazilian Portuguese and the Quantity Parameter. Gorka Elordieta and Marina Vigário (eds.) *Journal of Portuguese Linguistics Vol 5/6, Special Issue on the Prosody of the Iberian Languages:* 9–58.

Wetzels, W. Leo & Willebrord Sluyters. 1995. Formação de Raiz, Formação de Glide e 'Decrowding' Fonético em Maxacalí. Leo Wetzels (ed.) *Estudos Fonológicos de Língüas Indigenas Brasileiras.* Editora da UFRJ. Rio de Janeiro: 103–149.

Index of subjects and terms

For many different reasons, speakers borrow words from other languages to fill gaps in their own lexical inventory. The past ten years have been characterized by a great interest among phonologists in the issue of how the nativization of loanwords occurs. The general feeling is that loanword nativization provides a direct window for observing how acoustic cues are categorized in terms of the distinctive features relevant to the L1 phonological system as well as for studying L1 phonological processes in action and thus to the true synchronic phonology of L1. The collection of essays presented in this volume provides an overview of the complex issues phonologists face when investigating this phenomenon and, more generally, the ways in which unfamiliar sounds and sound sequences are adapted to converge with the native language's sound pattern. This book is of interest to theoretical phonologists as well as to linguists interested in language contact phenomena.

JOHN BENJAMINS PUBLISHING COMPANY