Citation: Padgett, Jaye (2010). Russian /Cju/ and "perceptual" vs. "phonological" theories of borrowing: a reply to Paradis (and Thibeault). Lingua 120.5, 1233-1239.

Russian /C^{j}u/ and "perceptual" vs. "phonological" theories of borrowing: a reply to Paradis (and Thibeault)

Jaye Padgett – UC Santa Cruz

1. Introduction

When Russian borrows a word having front rounded [y] in it, such as French *costume*, this non-native sound is normally rendered by means of [u] following a palatalized consonant: [kɐ'st^{j}um] 'suit, costume'. It is a question of some interest to phonologists why sounds get adapted and how precisely a given non-native sound will be adapted. Work in recent years has emphasized a competition between a "phonological" approach to these questions and a "perceptual" one.[1] According to the "phonological" view (e.g., Jacobs & Gussenhoven 1998; LaCharité & Paradis 2005; Paradis 1996), the underlying form of a borrowed word in L1 is equivalent to its (phonological) surface form in L2; adaptation happens when L1's phonology alters illicit structure found in this UR. For the case at hand, this would mean that *costume* was borrowed into Russian as /kɔ'stym/; this form's emergence as [kɐ'st^{j}um] followed from Russian phonological rules or constraints. This explains the reduction of /ɔ/ to [ɐ], a process amply present in Russian phonology. There is no independent evidence from Russian phonology for [y] → [^{j}u] (that is, [u] preceded by a palatalized sound), precisely because native Russian phonology lacks a phoneme [y]. But calling on a language's phonology to explain this sort of transformation is a common move in current phonological theory; such "repairs" are even required for example by Optimality Theory's tenet of richness of the base.

The "perceptual" approach to adaptation (e.g., Peperkamp & Dupoux 2003) differs in assuming that L1 speakers often *cannot* posit L2 surface forms as L1 underlying forms directly. Instead these L2 surface forms are first "filtered" according to the L1 speaker's perception. Intuitively speaking, L1 speakers "mishear" L2 surface forms, so that the URs they posit are already distorted from the L1 original. According to this view, adaptation of [y] in *costume* might happen because Russian borrowers *perceive* [y] as [^{j}u]. They therefore posit the UR /kɔ'st^{j}um/ to begin with.

It is well known that perception and categorization of speech is structured by the phonological categories of L1 (Best 1994). At the same time, some researchers have argued that phonological theory should incorporate explicitly perceptual notions (e.g., Flemming 1995 [2002]; Flemming 2003; Steriade 1997; Steriade 2001). Boersma (1998; 2000) argues that perception itself is under the control of the grammar, implementing this idea in the framework of

[1] My characterization of the difference basically follows (Paradis 2006). Many researchers combine perceptual and phonological adaptation in some way, e.g., Silverman (1992), Yip (2002).

Optimality Theory. Boersma and Hamann (2008) argue that both production and perception of loanwords are determined by the L1 grammar and can be modeled via mappings within Optimality Theory. For these reasons a terminological opposition between "perceptual" and "phonological" is at best unhelpful. From here on I will refer to the two approaches to loanword adaptation described at the outset as the *perception-* and *production-based* views. Though conceptually different, the perception- and production-based views are often difficult to distinguish empirically. I return to this point in the conclusion.

Paradis and Thibeault (2004) and Paradis (2006) argue that adaptation of [y] as [^{j}u] by Russian speakers must be specifically production-based. Their argument goes like this: a) the sequence [C^{j}u] is both phonetically unnatural and rare in Russian; b) it is impossible to misperceive a sound (sequence) as something that is unnatural and/or rare in your own language; c) therefore the rendering of [Cy] as [C^{j}u] must be production-based.[2]

This article is a brief reply to Paradis (2006) and Paradis and Thibeault (2004). Its main goal is to show that, contrary to (a), the sequence [C^{j}u] is both commonplace in Russian and natural. Whatever we think of the premise (b) therefore, the conclusion (c) does not follow. Some general discussion of the production- versus perception-based approach to adaptation follows, but neither view is argued for.[3]

2. Russian [C^{j}u] is not rare

The central premise of Paradis and Thibeault (2004) and Paradis (2006) is that [C^{j}u] is both unfamiliar to Russian speakers and phonetically unnatural. The claim of unfamiliarity or rarity is made repeatedly: "Generally speaking, Tables 1 and 2 indicate how infrequent and unnatural the sequence /C^{j}u/ is in Russian (p. 979)"; "the sequence /C^{j}u/, although phonetically unnatural, and unfamiliar in Russian... (p. 980)"; "the sequence /C^{j}u/ is clearly atypical in the Russian lexicon... (p. 992)" (all quotes from Paradis 2006). This claim is fundamental to the conclusion that a perception-based theory cannot handle the Russian facts: "By definition, 'perceptual deafness' cannot account for the sequence /C^{j}u/, which results from the adaptation of the foreign vowel /y/ after a palatalizable consonant in Russian, since this sequence is perceptually abnormal in this language (pp. 984-5)".

Paradis and Thibeault base the claim of rarity on a dictionary count of word-initial CV sequences, where C is any of [p,b,t,d]. They found that C is always palatalized before a front vowel (either of [i,ɛ]) but is palatalized only about 2% of the time before back vowels (any of [a, ɔ,u]); the figure is about 5% before [u] specifically.

---

[2] As one reviewer points out, this argument begs the question whether a production-based ("phonological") account should be deriving "rare" and "unnatural" sequences, and if so, why production is different in this way.

[3] There are a few other errors in Paradis and Thibeault (2004) and Paradis (2006) worth correcting: palatalized consonants are traditionally called "soft" (Russian *m^{j}agkij*), not "wet" (Russian *mokrij*); /a/ reduces to [ɐ/ə], not [i], unless a soft consonant comes before; [Cju] (with a glide, not palatalization) does in fact occur in Russian, and is distinct from [C^{j}u] (see section 4).

It is not obvious that the frequency of [C^{j}u] relative to [C^{j}i/C^{j}ɛ] is the right criterion to use. But putting this aside, the choice to focus on word-initial CV for C = [p,b,t,d] was unfortunate in a couple of respects. First, a rich source of native [C^{j}u] in Russian is found at stem-inflection boundaries. For example, [C^{j}u] occurs when various kinds of verb are put in the first person singular, as shown in (1).[4] These words are as routine as their glosses suggest. To make reading transcriptions easier, for the rest of the paper I abstract away from vowel reduction (which never affects /u/) and from certain other surface effects, such as assimilation of palatalization.

(1) [C^{j}u] in 1st person sg.

1st person sg.	gloss	*cf.*	infinitive
[l^{j}u'bl^{j}u]	'I love'		[l^{j}u'b^{j}it^{j}]
[xɔ'$\widehat{\text{tʃ}}^{j}$u]	'I want '		[xɔ't^{j}it^{j}]
[gɔvɔ'r^{j}u]	'I speak, say'		[gɔvɔ'r^{j}it^{j}]
['pɔmn^{j}u]	'I remember'		['pɔmn^{j}it^{j}]
[i'ʃ^{j}:u][5]	'I look for'		[i'skat^{j}]

[C^{j}u] also occurs whenever a noun which takes the nominative ending /-a/ (mostly feminine), having a stem ending in a palatalized consonant, is put in the accusative case:

(2) [C^{j}u] in acc. sg. of nouns in /-a/ (C^{j}-final stem)

acc. sg.	gloss	*cf.*	nominative sg.
['d^{j}ad^{j}u]	'uncle'		['d^{j}ad^{j}a]
[n^{j}ɛ'd^{j}ɛl^{j}u]	'week'		[n^{j}ɛ'd^{j}ɛl^{j}a]
['kat^{j}u]	'Katya (acc.) ' (proper name)		['kat^{j}a]
['kuxn^{j}u]	'kitchen'		['kuxn^{j}a]
['bur^{j}u]	'storm'		['bur^{j}a]

[4] This occurs with verb stems ending in either [i], [ɛ], or [ɔ]; the specific sequences [$\widehat{\text{tʃ}}^{j}$u] and [ʃ^{j}:u] occur also when verb stems end in [$\widehat{\text{tʃ}}^{j}$a] or [ʃ^{j}:a], where the consonant is either underlying or derived by mutation. In [ɔ]-stem verbs and in [a]-stem verbs with mutation, [C^{j}u] also occurs in the third person plural, e.g., ['bɔr^{j}uts^{j}a] 'they fight' and [bɔr'mɔ$\widehat{\text{tʃ}}^{j}$ut] 'they mutter', cf. [bɔ'rɔts^{j}a] 'to fight' and [bɔrmɔ'tat^{j}] 'to mutter'. Verbs in [i] are productive and number in the thousands according to Townsend (1968).

[5] The sound [ʃ^{j}:] of Contemporary Standard Russian of Moscow, often transcribed as [ɕ:] or [ɕ'č'] by slavists, is historically derived from palatalized clusters like [st^{j}]; hence the length, which is reliably preserved at least intervocalically. In some dialects it is still a cluster [ʃ$^{j}\widehat{\text{tʃ}}^{j}$].

One more inflectional context leading to [C^{j}u] is that of C^{j}-final masculine or neuter stems in the dative case:

(3) [C^{j}u] in dat. sg. of masculine C^{j}-final stem nouns

dat. sg.	gloss	*cf.*	nominative sg.
['mɔr^{j}u]	'sea'		['mɔr^{j}ɔ]
['pɔl^{j}u]	'ground'		['pɔl^{j}ɔ]
['dn^{j}u]	'day'		['d^{j}ɛn^{j}]
['gɔst^{j}u]	'guest'		['gɔst^{j}]
[vra'$\widehat{\text{tʃ}}^{j}$u]	'doctor'		['vrat$\widehat{\text{ʃ}}^{j}$]

The second reason why focusing on word-initial [p,b,t,d] underrepresents [C^{j}u] is that the most frequent examples of [C^{j}u] involve consonants other than [p,b,t,d]. The reason is historical.[6] Many or most instances of [C^{j}u] within Slavic roots are derived from Proto-Slavic [Cju] (with a glide). A Proto-Slavic sound change called 'jotation' mutated consonants standing before [j] to palatalized C^{j}. When the following vowel happened to be [u] the result was [C^{j}u]. (The Proto-Slavic sequence [Ceu] appears to have undergone parallel development.) But various concomitant or subsequent changes, entirely independent of any following [u], narrowed the range of consonants to be found in [C^{j}u]: with labials jotation was accompanied by insertion of [l] (compare the forms of 'love' in (1)); in addition, both velar and dental obstruents mutated to palato-alveolar place, and some of those palato-alveolars later depalatalized. A result of these changes is that [l^{j},r^{j},n^{j},$\widehat{\text{tʃ}}^{j}$,ʃ^{j}:] are the consonants found most typically in [C^{j}u]. (The verbs in (1) are representative.)

But a number of independent factors, including borrowings, paradigm uniformity effects, and lexical innovations (all of which led to [C^{j}u] already in the earliest stages of Russian, many hundreds of years ago), have extended the range of consonants found in [C^{j}u] to include virtually every palatalizable consonant in Russian, including relatively frequent forms having [b^{j},t^{j},d^{j},s^{j}].[7] The forms shown in (4), everyday Russian words, are from a list of the 5,000 most frequent lemmas found in a corpus of about 40 million words (Sharoff 2008).[8] Those in (4)a are Slavic forms, many hundreds of years in use; those in (4)b, though historically borrowed, are core

---

[6] For discussion of these facts see Carlton (1991) and Townsend and Janda (1996).

[7] Apart from [l^{j},r^{j},n^{j},$\widehat{\text{tʃ}}^{j}$,ʃ^{j}:], these are the consonants most frequently found in [C^{j}u] sequences. This is according to a corpus of 32,000 of the most frequent forms (occurring more than once per million words) occurring in Sharoff (2008) (see below).

[8] According to Sharoff (2008), the corpus "contains selection [sic] of modern fiction, political texts, newspapers, and popular science (about 40 million words…fiction allocates for about half of the corpus). All texts were written originally in Russian between 1970 and 2002; the majority of them between 1980 and 1995, the newspapers [sic] corpus is from 1997-1999."

Russian vocabulary items. Some such forms (e.g., [tʲur'ma] ‘prison’) are also many hundreds of years old.

(4) A selection of lemmas from a Russian corpus (Sharoff 2008)

a.	lʲu'bɔvʲ	‘love’	b.	kɔ'stʲum	‘suit, costume’
	sʲu'da	‘hither’		tʲurʲ'ma	‘prison’
	ɔt'nʲudʲ	‘by no means’		'brʲukʲi	‘trousers’
	'batʲuʂka	‘father’		bʲu'dzɛt	‘budget’
	'ʧʲutʲ	‘hardly’			
	ɔʃʲ:u'ʃʲ:atʲ	‘feel, sense’			

Within this list of the 5,000 most frequently occurring Russian lemmas, there are 79 forms containing [Cʲu]. Though [Cʲu] in lemmas (to be distinguished from inflectional endings) might not win a prize for frequency, even there [Cʲu] is very familiar to Russian speakers.

One more fact worth mentioning underscores the commonplace status of [Cʲu] in Russian. The language occasionally “spontaneously” produces new words having palatalized consonants before [u]; this is thought to occur in at least some cases under the influence of related or similar-sounding words having palatalization. The core vocabulary word [sʲu'da] ‘hither’, for example, seems to have arisen from an Old Russian form [su'da]. The word ['dʲuʑina] ‘dozen’ comes from French *douzaine* [du'zɛn]; the palatalization is speculated to have arisen under the influence of ['dʲuʑij] ‘hefty, strapping’, itself a native word thought to derive from a form ['duʑij] (see Vasmer 1964-1973 for these and other cases.).

3. [Cʲu] is not unnatural

Paradis and Thibeault (2004) and Paradis (2006) are equally forceful in claiming that the sequence [Cʲu] is unnatural. To the extent that “unnatural” means “unfamiliar to Russians” this is the claim discussed above. But the idea also appears to be that [Cʲu] is phonetically unnatural, based on the observation that assimilatory palatalization happens only before front vowels, never before back vowels, e.g., “back vowels do not naturally yield consonant palatalization” (Paradis 2006:978). Paradis (2006) also claims that there are languages in which consonants are palatalized before front vowels but not back vowels, but no languages having the reverse distribution.

It is true that consonants palatalize by assimilation only before front vocoids. But it follows that [Cʲu] is unnatural only if one assumes that the only bestower of “naturalness” is assimilation. This cannot be true. Languages in fact place a high value on sequences that *disagree* in features; the most fundamental indication of this is the tendency for consonants to alternate with vowels. When palatalization is *contrastive* and one considers the perceptual side of naturalness, no vowel could be better after [Cʲ] than [u] (compare [bʲu] vs. [bu] with [bʲi] vs. [bi]), and the implicational claim above is turned on its head: a language is more likely to *neutralize* a palatalization contrast – sometimes to *unpalatalized* consonants – before front vowels (Ní Chiosáin & Padgett 2001; Padgett 2001). One might compare the tendency for languages to avoid sequences like [ji,je] –

when in *contrast* with [i,e] – in favor of [ja,jo,ju] (contrastive with [a,o,u]). Presumably [ju] is not unnatural.

4. Discussion and conclusion

If Russian [C^{j}u] is both familiar to Russian speakers and natural, then the Russian adaptation of [Cy] as [C^{j}u] is not (at least so far) an argument for the production-based view of loanword adaptation. It is equally compatible with a perception-based view.

Below is an illustrative production-based analysis of the Russian adaptation within the framework of Optimality Theory, one which basically carries over essential aspects of the analysis of Paradis and Thibeault (2004) and Paradis (2006). The claims of such an analysis are that Russian phonology a) prohibits front rounded vowels; b) prefers to repair such vowels by altering their backness, not their roundness; and c) requires that the underlying [-back] value be retained (cf. the Preservation Principle of Paradis & LaCharité 1997). I assume that other constraints not shown require that [t] be palatalized before a front vowel, that /ɔ/ be reduced, and that other outcomes which preserve [-back], such as [kɐ'stuim] be avoided.

(5) A production-based analysis of /Cy/ → [C^{j}u] adaptation in Russian

	/kɔ'stym/	*[-bk, +rd]	Ident[round]	Max[back]	Ident[back]
a.	kɐ'stym	*!			
b.	kɐ'st^{j}ym	*!			
c. ☞	kɐ'st^{j}um				*
d.	kɐ'st^{j}im		*!		
e.	kɐ'stum			*!	*

This sort of analysis seems very reasonable. But a perception-based account is now also reasonable, given what we saw in the last two sections: a Russian listener *perceives* [Cy] as the sequence [C^{j}u]; in more detail, they find [C^{j}u] to be the best match for [Cy], given Russian perceptual categories. The two accounts clearly have a great deal in common: a constraint *[-bk, +rd] substitutes for a perception-based account's observation that Russian has no category "front rounded vowel"; the deployment of faithfulness constraints and rankings shown substitutes for the claim that [Cy] is perceptually closest to [C^{j}u].

Paradis and Thibeault (2004) and Paradis (2006) see one final argument for the production-based approach in the following fact: after consonants that cannot be palatalized in Russian, as well as word-initially, [y] is adapted as simple [u]. According to these authors, under a perception-based account we might instead expect [ju], since [ju] is licit in these contexts and

presumably is a closer perceptual match to [y].[9] But this does not happen, e.g., [zju'ri] (*[zju'ri]) 'jury, judges' < French *jury*, and [unʲi'tarnʲij] (*[junʲi'tarnʲij]) < French *unitaire* (the latter example from Paradis and Thibeault).

This last argument seems to depend on an inference: if [Cy] is more similar to [Cʲu] than to [Cu], then [y] should be more similar to [ju] than to [u]. But the two cases are not parallel in Russian, either phonetically or phonologically. Consider the hypothetical scenario in (6)a, in which a Russian listener perceives French [by] and attempts to categorize it. The relevant choices in Russian (assuming solutions with [i] are ruled out) are /bu/, /bʲu/, and /bju/. The difference between /Cʲu/ and /Cju/ is no triviality in Russian; in fact it is a contrastive difference, as in the near-minimal pair ['volʲu] 'will (acc.sg.)' and ['bolju] 'pain (instr.sg.)'. The Russian listener has three categorically distinct forms to choose from. If a perceptual approach to adaptation is right, then [by] must fall most closely within the category space of /bʲu/, as indicated in the diagram. The situation when a Russian perceives [y], in a context where palatalization is impossible, is different, as shown in (6)b. It should be clear that (6)a gives us no basis to assume that /ju/ is the closest match in (6)b. The choice between /ju/ and /u/ is also not obvious on phonetic grounds; /u/ would not be associated with the higher second formant of front [y], but [y] lacks the very robust transition from high to low F2 that would signal /ju/.

(6) Hypothetical perceptual categorizations of (C)y in Russian without palatalization

The case of Russian [y] → [ʲu] is typical of many adaptations, in that both perception- and production-based theories seem to offer reasonable accounts of the facts; it can be difficult to distinguish empirically between the two approaches.[10] They are both phonological, in the sense that they assume that their respective processes are structured by phonological categories and are context dependent. In addition phonological theory, as practiced by some, is taking on increasing phonetic knowledge such as access to phonetic cues, or even modelling perception itself, further narrowing the distance between "phonological" and "perceptual" accounts.

---

[9] The non-palatalizable consonants (ignoring [j]) are [t͡s,ʂ,ʐ]. Paradis and Thibeault include the velars [k,g,x] as well. The latter are *allophonically* palatalized before front vowels, but Paradis and Thibeault have in mind the back vowel context. Even there the status of [k,g,x] as non-palatalizable is debatable (see Padgett 2003:46-7): they do occur palatalized in loanwords (e.g., [manʲi'kʲur] 'manicure'). Paradis and Thibeault's argument cannot be extended to velars anyway, because [ju] is not in fact licit after velars.

[10] A potential argument for the perception-based approach in some cases is that loanword adaptation does not mirror how illicit structures are repaired in the L1 phonology (Boersma and Hamann 2008). The alternative is to handle such cases by means of loanword-specific rankings or constraints.

Perception-based accounts have the potential to make falsifiable predictions that can be experimentally tested, though the debate about adaptation has benefitted little from this possibility so far. (However, see Peperkamp et al. 2008 and references therein.) One might likewise ask whether proponents of production-based approaches can better articulate – for cases like Russian – testable predictions distinguishing production-based accounts from perception-based ones.

Acknowledgements

I am very grateful to David Teeple, Abby Kaplan, and three anonymous reviewers for comments on this paper, and to Alex Matthews for help with Perl scripting.

References

Best, C. T. 1994. The emergence of native-language phonological influences in infants: a perceptual assimilation model. In: Nusbaum, H. C. & J. Goodman (eds.), The development of speech perception: the transition from speech sounds to spoken words. Cambridge, MA, MIT Press, 167-224.

Boersma, P., 1998. Functional phonology: formalizing the interactions between articulatory and perceptual drives. The Hague, Holland Academic Graphics.

Boersma, P., 2000. The OCP in the perception grammar. Rutgers Optimality Archive 435.

Boersma, P. & S. Hamann, 2008. Loanword adaptation as first-language phonological perception. Ms., University of Amsterdam and Universität Düsseldorf.

Carlton, T. R., 1991. Introduction to the phonological history of the Slavic languages. Columbus, OH, Slavica Publishers.

Flemming, E., 1995 [2002]. Auditory representations in phonology. Doctoral dissertation, UCLA [published in 2002 by Routledge, New York].

Flemming, E., 2003. Speech perception in phonology. Ms., Stanford University.

Jacobs, H. & C. Gussenhoven. 1998. Loan phonology: perception, salience, the lexicon and OT. In: Dekkers, J., F. van der Leeuw & J. van de Weijer (eds.), Optimality Theory: phonology, syntax, acquisition. Oxford, Oxford University Press, 193-210.

LaCharité, D. & C. Paradis, 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. Linguistic inquiry 36.2, 223-258.

Ní Chiosáin, M. & J. Padgett. 2001. Markedness, segment realization, and locality in spreading. In: Lombardi, L. (ed.) Segmental phonology in Optimality Theory: constraints and representations. Cambridge, Cambridge University Press, 118-156.

Padgett, J. 2001. Contrast dispersion and Russian palatalization. In: Hume, E. & K. Johnson (eds.), The role of speech perception in phonology. San Diego, CA, Academic Press, 187-218.

Padgett, J., 2003. Contrast and post-velar fronting in Russian. Natural language and linguistic theory 21.1, 39-87.

Paradis, C. 1996. The inadequacy of filters and faithfulness in loanword adaptation. In: Durand, J. & B. Laks (eds.), Current trends in phonology: models and methods. Salford, University of Salford, 509-534.

Paradis, C., 2006. The unnatural /C^{j}u/ (< foreign /Cy/) sequence in Russian loanwords: a problem for the perceptual view. Lingua 116, 976-995.

Paradis, C. & D. LaCharité, 1997. Preservation and minimality in loanword adaptation. Journal of linguistics 33, 379-430.
Paradis, C. & M. Thibeault, 2004. L'adaptation de la voyelle /*y*/ en /^{j}*u*/ dans les emprunts en russe: un cas de palatalisation non de biphonémicité. Canadian Journal of Linguistics 49.2, 175-222.
Peperkamp, S. & E. Dupoux. 2003. Reinterpreting loanword adaptations: the role of perception. Barcelona, Proceedings of ICPhS 15, 367-370.
Peperkamp, S., I. Vendelin & K. Nakamura, 2008. On the perceptual origin of loanword adaptations: experimental evidence from Japanese. Phonology 25, 129-164.
Sharoff, S., 2008. The frequency dictionary for Russian. Moscow, Russian Research Institute of Artificial Intelligence.
Silverman, D., 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. Phonology 9, 289-328.
Steriade, D., 1997. Phonetics in phonology: the case of laryngeal neutralization. Ms., UCLA.
Steriade, D. 2001. Directional asymmetries in place assimilation: a perceptual account. In: Hume, E. & K. Johnson (eds.), The role of speech perception in phonology. New York, Academic Press, 219-250.
Townsend, C. E., 1968. Russian word-formation. New York, McGraw-Hill.
Townsend, C. E. & L. A. Janda, 1996. Common and comparative Slavic: phonology and inflection. Columbus, Ohio, Slavica.
Vasmer, M., 1964-1973. Etimologicheskii slovar' russkogo iazyka. Moscow, Progress.
Yip, M., 2002. Perceptual influences in Cantonese loanword phonology. Journal of the Phonetic Society of Japan 6, 4-21.