Testing Licensing by Cue: A case of Russian palatalized coronals

Alexei Kochetov, University of Toronto

Available at: https://works.bepress.com/alexei_kochetov/11/
Abstract
The hypothesis ‘licensing by cue’ by Steriade holds that phonological contrasts are maintained in environments that provide better acoustic cues to the contrasts and are neutralized in environments that provide poorer acoustic cues or no cues. This paper tests the hypothesis by examining the distribution of a phonological contrast – the Russian plain/palatalized coronal stops /t/ and /tʲ/ in various syllable-final contexts. The results of a series of acoustic and perceptual experiments presented in this paper provide some support for the hypothesis: the relative salience of releases in different word boundary contexts (_#k > _#n, _#s) correlates strongly with the general patterns of neutralization of the contrast in similar word-internal contexts (_k > _n, _s) in Russian and other related languages. At the same time, the relative salience of VC transitions in different vowel contexts (a_ > u_ > i_) has apparently little to do with attested patterns of neutralization. The results suggest that some perceptual cues are phonologically more relevant than others, providing evidence for interactions between phonetics and phonology more complex than predicted by the hypothesis.

1 Introduction

licensing by cue
The hypothesis of licensing by cue makes straightforward predictions about the relationship between the acoustics and perception of a particular phonological contrast in a particular phonotactic environment and the presence or absence of neutralization in that environment. If we establish that among two contexts A and B, context A provides more acoustic information about some phonological contrast x versus y than context B does, this implies that the contrast is more perceptually recoverable in context A than in context B. Ultimately, this implies that the contrast is more likely to be maintained in context A (x vs. y\textsubscript{A}) and more likely to be neutralized in context B (*x vs. y\textsubscript{B}), not the other way around. We would, therefore, expect to find languages that maintain the contrast in context A and neutralize it in context B.

The goal of this paper was to test the licensing by cue hypothesis by examining the role of acoustic cues in the complex distribution of plain and palatalized denti-alveolar stops /t/ and /t\textsuperscript{Js} in Russian (e.g. [t\textbar{omni\j}] 'obese' vs. [t\textbar{omni\j} \textsuperscript{Js}] 'dark', [ma\textbar{t}] 'foul language' vs. [ma\textbar{t}\textsuperscript{Js}] 'mother'). Importantly, the contrast /t/ versus /t\textsuperscript{Js} is maintained in certain word-internal contexts and neutralized in others; it is always maintained in word-final contexts, both utterance finally and at word boundaries. To determine whether the maintenance and neutralization of the contrast word-internally is phonetically motivated, we will examine acoustic and perceptual properties of the contrast in similar word boundary clusters. Further, to determine the relevance of phonetic information and the maintenance of the contrast after different vowels and word-finally, we will examine acoustic and perceptual properties of the contrast in a number of utterance-final contexts. Before turning to the experiments, we will review the acoustics and phonological distribution of the contrast /t/ and /t\textsuperscript{Js} in Russian.

2 The Acoustics of Russian Plain and Palatalized Coronal Stops

2.1 VC and CV Transitions
This effect was attributed to the fronting and raising of the front of the tongue to the palate (as well as the pharyngeal widening) simultaneously with the primary articulation. Based on an area function derived from X-ray images, Fant [1970, p. 223] predicted that 'terminal' F2 values (loci) associated with plain and palatalized denti-alveolars should differ by about 530 Hz (1,370 vs. 1,900 Hz). His acoustic measurements of F2 of these consonants at the onset of the following vowel /a/ revealed a smaller difference of 240 Hz (1,700 vs. 1,440 Hz for a male speaker; 1,700 vs. 1,300 Hz in Bondarko [1977, pp. 146–147]).

In most vowel contexts, palatalized consonants were noted to be characterized by more 'positive' F2 transitions (rising VC and falling CV transitions) than their plain counterparts; in the contexts of the front vowels /i/ and /ε/ palatalized consonants showed 'zero' (level) F2 transitions, while plain consonants showed 'negative' F2 transitions (falling VC and rising CV transitions) [Halle 1959/1971, pp. 152–153; Purcell, 1979; see also Padgett, 2001, on CV transitions in the /i/ context]. Differences between vowel contexts in transitions to palatalized consonants appear to have considerable perceptual consequences, at least for syllable-final palatalized consonants. Positive transitions seem to be better cues to palatalization than zero transitions. According to Jones and Ward [1969, pp. 94, 104], the [i]-like on-glide typically heard prior to syllable-final palatalized consonants is 'barely discernible' or 'not discernible at all' in the context of front vowels.

2.2 Releases

An important aspect of the production of Russian syllable-final (word-final and preconsonantal) stops is their audible releases [Jones and Ward, 1969]. Thus, Russian utterance-final stops are always audibly released, with the releases accompanied by strongly aspirated either velarized or palatalized off-glides [Bondarko, 1981, pp. 104, 132–133; Fant, 1970, p. 223]. Releases are somewhat less frequent when word-final consonants are followed by other consonants. Zsiga [2000] found that the stops /p, d, k/ in Russian word boundary sequences (C1 in C1#C2) were released on average half of the time, with a lower percentage of releases in homorganic clusters (e.g. less than 10% for /d#t/) and a higher percentage of releases in hetero-organic clusters (e.g. 100% for /d#k/). This finding confirms the auditory observations of Jones and Ward [1969, pp. 88, 89, 105] that Russian stops are often released in hetero-organic clusters and are unreleased in homorganic clusters (or have nasal or lateral releases when followed by hetero-organic nasals or laterals). Results of an articulatory EMMA study [Kochetov and Goldstein, 2001] showed that the gestures of Russian stops in word boundary hetero-organic clusters were in fact produced with a considerable lag (no overlap) allowing for a release of C1. The same stops in homorganic clusters often lacked articulatory releases, forming one long constriction together with C2.

Examining spectral properties of syllable-initial plain and palatalized stops, Halle [1959/1971, p. 150] noted that the most striking difference between /t, d/ and /t, d/ was a considerable amount of affrication in the latter [Jones and Ward, 1969, pp. 104, 107; Bondarko, 1981, p. 111]. His examination of burst spectra, however, did not reveal consistent differences between plain and palatalized coronals (p. 152). Fant [1970] described syllable-initial /t/ and /t/ as both having energy between 500 and 8,000 Hz.
2.3 Summary

Previous research has shown that the contrast between the Russian /t/ and /tʃ/ is cued by a number of acoustic parameters, primarily by (a) F2 of the following and preceding vowels and (b) duration, amplitude, and (less consistently) spectra of the release (burst and frication). The two sets of acoustic parameters will be referred to as (a) transitions and (b) releases. To our knowledge, none of the previous studies have consistently examined transitions and releases of /t/ and /tʃ/ in syllable-final position (word-internally or word-finally across word boundaries and utterance finally). While the perceptual importance of CV transitions as a cue to the plain/palatalized contrast has previously been documented [Derkach, 1975; Kochetov, 2004; Kavitskaya, in press], no studies have focused on the role of and relative perceptual importance of VC transitions and releases. Examining cues to the voicing contrast, Steriade [1997, pp. 5, 6] argued that stop releases ('onset cues' such as VOT) have perceptual primacy over other cues to voicing ('offset cues' and 'internal cues' such as V1 duration, closure duration, F1 or F0 values). It is tempting to extend this argument to the plain/palatalized contrast in stops, especially in the light of previous perceptual studies that confirmed the important role of releases in the identification of stop place [e.g. Malécot, 1958]. At least, one might expect releases to be more salient than VC transitions, which are notably perceptually inferior to CV transitions (see e.g. Ohala [1990] on place and Kochetov [2004] on palatalization).

The goal of the experiments presented in this paper was to determine how VC transitions and releases differentiate Russian /t/ and /tʃ/ in a number of syllable-final contexts. This will help us to answer the key question whether the distribution of acoustic cues to the contrast is relevant to the phonotactic distribution of the contrast, which is discussed in the next section.
### 3 The Distribution of Russian Plain and Palatalized Coronal Stops

<table>
<thead>
<tr>
<th>Plain</th>
<th>Palatalized</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>t</strong></td>
<td>'so'</td>
<td>drought</td>
</tr>
<tr>
<td>'tomnij'</td>
<td>'obese'</td>
<td>dark</td>
</tr>
<tr>
<td>'turk'</td>
<td>'rook'</td>
<td></td>
</tr>
<tr>
<td>test</td>
<td>'test'</td>
<td>dough</td>
</tr>
<tr>
<td>'tič'</td>
<td>'to poke'</td>
<td>to tick</td>
</tr>
<tr>
<td>t</td>
<td>'foul language'</td>
<td>mother</td>
</tr>
<tr>
<td>ot</td>
<td>'walk'</td>
<td>though</td>
</tr>
<tr>
<td>t</td>
<td>'blown'</td>
<td>to blow</td>
</tr>
<tr>
<td>'t'</td>
<td>'beaten'</td>
<td>to beat</td>
</tr>
<tr>
<td>'č'</td>
<td>'sung'</td>
<td>to sing</td>
</tr>
</tbody>
</table>

### 4 The Distribution of Russian Plain and Palatalized Coronal Stops

- **Plain** /t/ and /t/ are contrastive in syllable-initial position before /k/ and /l/.
- In other syllable-final contexts, the contrast is maintained or neutralized depending on the nature of the following consonant.
- Before hetero-organic context before homorganic consonants, Russian stops tend to lack audible releases.
- In syllable-initial position, the phonological quality—backness and height—is maintained regardless of the following or preceding vowel.
- The contrast between /t/ and /t/ is maintained regardless of the phonological quality.
- Testing Licensing by Cue
- See also Farina [1991] for an underspecification account.
voicing contrast is neutralized to voiceless stops in word-final position and before voiceless obstruents. The contrast between articulation [Jones and Ward, 1969; Avanesov, 1972]. Thus, plain and palatalized labial and dorsal stops /p, pʃ/. While the focus of the paper is on the contrast between the anterior coronal stops /t, tʃ/ and /t, tʃ/ in word-internal derived environments (as in 3a). The contrast between plain and palatalized coronal affricates /ts, tʃ/ is sensitive to the place of articulation (with /p, pʃ/ being neutralized to /p/). Further, all voiceless stops have voiced counterparts /d, dʒ/ as well as the plain and palatalized coronal affricates /tʃ, tʃ/.

Turning to the hypothesis of licensing by cue, the maintenance of the contrast and its neutralization are directly related to the amount of acoustic information potentially available from the preceding and following consonants (e.g. _k) presumably because the contrast in these contexts is perceptually cast as a plain/palatalized contrast.

It should be noted that the depalatalization of /tʃ/ before homorganic consonants is sensitive to the quality of the preceding vowel. Importantly, whether preceding vowels are back or front (/u, o, a/ vs. /i, e) does not play any role in the maintenance (1 and 2) of the contrast, whether the following consonant is hetero-organic (4a) or homorganic (4b). [Jones and Ward, 1969; Avanesov, 1972]

Further, all voiceless stops have voiced counterparts /d, dʒ/ as well as the plain and palatalized coronal affricates /tʃ, tʃ/.

To summarize, the contrast between /t/ and /tʃ/ before hetero-organic consonants (coronals). The contrast is always maintained across word boundaries. Further, whether the following consonant is hetero-organic (4a) or homorganic (4b). Within words, it is maintained before hetero-organic consonants (velars and labials), and neutralized to the plain /t/ before homorganic consonants.

Further, the contrast is not sensitive to the quality of the preceding vowel. It is important to mention that the described distributional pattern of standard Russian (Belorussian, Ukrainian, and Bulgarian) has historically developed their patterns of neutralization independently from Russian [Carlton, 1990]. Further, in none of these languages does the qualitative pattern with no neutralization altogether (an eastern Bulgarian dialect; see Kochetov 2002 for a language survey). Importantly, at least some of these languages (Polish and Bulgarian) have historically developed their patterns of neutralization independently from Russian [Carlton, 1990]. Other Slavic languages with a similar contrast show whether the following consonant is hetero-organic (4a) or homorganic (4b). The only apparent deviation involves additional neutralization of the contrast before hetero-organic stops and word-finally in certain northern Russian dialects [Kuznetsova, 1969]. Other Slavic languages with a similar contrast show the contrast before hetero-organic stops and word-finally in certain northern Russian dialects [Kuznetsova, 1969]. The only apparent deviation involves additional neutralization of the contrast before hetero-organic stops and word-finally in certain northern Russian dialects [Kuznetsova, 1969].

It is important to mention that the described distributional pattern of standard Russian (Belorussian, Ukrainian, and Bulgarian) has historically developed their patterns of neutralization independently from Russian [Carlton, 1990]. Further, in none of these languages does the qualitative pattern with no neutralization altogether (an eastern Bulgarian dialect; see Kochetov 2002 for a language survey). Importantly, at least some of these languages (Polish and Bulgarian) have historically developed their patterns of neutralization independently from Russian [Carlton, 1990]. Other Slavic languages with a similar contrast show whether the following consonant is hetero-organic (4a) or homorganic (4b). The only apparent deviation involves additional neutralization of the contrast before hetero-organic stops and word-finally in certain northern Russian dialects [Kuznetsova, 1969]. Other Slavic languages with a similar contrast show whether the following consonant is hetero-organic (4a) or homorganic (4b). The only apparent deviation involves additional neutralization of the contrast before hetero-organic stops and word-finally in certain northern Russian dialects [Kuznetsova, 1969].
4 Experiment 1: Production of Word-Internal Coronal Stops before /k/, /n/, and /s/

To summarize, the contexts of interest and relevant predictions are given in table 1, where word-internal and word-final contexts are given separately. The cells shown in bold correspond to the contexts directly investigated in this study. Thus, experiments 1 and 2 investigate the word-internal contexts a_k, as well a_n and a_s (for /t/ only) (table 1a); experiments 3 and 4 investigate the word-final contexts before consonants, a_#k, a_#n, and a_#s (table 1b), and experiments 5 and 6 investigate the utterance-final contexts a_#, u_#, and i_# (table 1b). The results of these experiments will also allow us to make predictions about the relevance of vowel quality in other word-final (_#n and _#s) and word-internal contexts (_k, _n, and _s).
4.1 Method

\[ \lambda'z\text{it}'\text{t}j'\text{f}:\text{c},\text{ras} \]
The latter two measurements were made to control for consonant-to-consonant coarticulation in /tk/ sequences. Acoustical and spectral measurements were made:

- Preceding vowel F2 frequency
  - $F_{2\text{mid}}$
  - $F_{2\text{off}}$
- Percent released (% released)
- Stop closure release duration
- Amplitude difference
- Stop release peak frequency
- Following stop closure duration (C2 closure)
- Following vowel F2 frequency
  - $F_{2\text{on}}$
  - $F_{2\text{off}}$

*consonant*;
*context*;
*speaker*

Recorded speech was digitized at a sampling rate of 22 kHz with 16-bit resolution, and analyzed acoustically using Praat's formant tracking with maximum formant values set at 5,000 Hz for male and 5,500 Hz for female speakers.

The collected data, averaged for two repetitions, were analyzed in an analysis of variance (ANOVA) with the independent variables speaker (_s); for /t/ only). The dependent variables were the acoustic measurements described above. Three-level speaker-specific deviations from the general trend (based on interspeaker variation) will be mentioned when relevant.

Whenever results differed between the two, the respective tokens were discarded. When determining audible perceptibility of releases, only a part of the signal was played: from the onset of the preceding vowel to the approximate midpoint of the closure of the following consonant. In the contexts of /n/ and /s/, the spectrogram contained a release burst [Zsiga, 2000], and (b) this release was perceptible auditorily [Henderson and Repp, 1982]. To minimize the bias in the identification of releases, this procedure was performed independently by the author and another trained phonetician.

The total number of recorded target words obtained was 288 (16 words per speaker). Only the second and third productions of each utterance were analyzed, given the relatively limited interspeaker variation, these statistical results will not be presented in detail. Significant speaker-specific deviations from the general trend (based on interspeaker variation) will be mentioned when relevant.

Testing Licensing by Cue

There was only 1 such token in experiment 1 and 4 tokens in experiment 3.

5 The measurements were performed using Praat's formant tracking with maximum formant values set at 5,000 Hz 6 Whenever results differed between the two, the respective tokens were discarded.7

Following vowel F2 frequency

- $F_{2\text{on}}$
- $F_{2\text{off}}$

Following stop closure duration (C2 closure)

- $F_{2\text{on}}$
- $F_{2\text{off}}$

Percent released (% released)
4.2 Results

Table 3 shows means and standard deviations for /t/ and /t/jsuperscript/ before /k/, as well as for /t/ before /n/ and /s/. ANOVA results are presented below, separately for the factors consonant and context. (Here and further, nonsignificant interactions of factors are not reported.) For reference, spectrograms of two sample tokens of /t/ and /t/jsuperscript/ before /k/ are given in the Appendix.

4.2.1 Consonant

With respect to the context before /k/, F 2 of the vowel was significantly higher before /t/jsuperscript/ than before /t/ at both points in time ([F(1, 71) = 11.171, p = 0.01; F(1, 71) = 76.022, p = 0.001; see table 3]. Both /t/ and /t/jsuperscript/ were always audibly released before /k/. The stops differed significantly in their duration: while /t/ had a longer closure and a shorter release, /t/jsuperscript/ had a shorter closure and a longer release ([F(1, 71) = 45.282, p = 0.001; F(1, 71) = 67.099, p = 0.001]). The release of /t/jsuperscript/ showed a smaller amplitude difference than the release of /t/ ([F(1, 71) = 21.248, p = 0.001], indicating that the release of /t/jsuperscript/ was louder than the release of /t/. The maximum amplitude for /t/ tended to occur at its release burst, while the maximum amplitude for /t/jsuperscript/ tended to occur later, during its fricative interval. Peak frequency of /t/jsuperscript/ measured at point 1 (at the burst) was higher than peak frequency of /t/ ([F(1, 71) = 62.532, p = 0.001]). Overall, the spectrum of /t/ was relatively ‘diffuse’ [Jakobson et al., 1963], showing high energy ranging from low frequencies up to 8,000 Hz. Given the relatively short release of /t/, in most cases it was not possible to measure its peak frequency at point 2. The spectrum of /t/jsuperscript/ was more ‘compact’, showing low energy below 4,000 Hz and high energy between 4,000 and 8,000 Hz, peaking around 5,500 and 6,000 Hz. Peak frequency of /t/jsuperscript/ was somewhat lower at point 2 (during the fricative period). The closure of the following consonant /k/ was slightly longer after /t/jsuperscript/ than after /t/ ([F(1, 71) = 6.999, p = 0.05]. F2 frequency at the offset of /k/ (at the onset of the following vowel /a/) was higher after /t/jsuperscript/ than after /t/ ([F(1, 71) = 6.601, p = 0.05].

### Table 3.

<table>
<thead>
<tr>
<th></th>
<th>a_k</th>
<th>a_n</th>
<th>a_s</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2 mid, Hz</td>
<td>1,499 (211)</td>
<td>1,652 (177)</td>
<td>1,328 (183)</td>
</tr>
<tr>
<td>F2 off, Hz</td>
<td>1,419 (197)</td>
<td>1,885 (134)</td>
<td>1,406 (225)</td>
</tr>
<tr>
<td>% released</td>
<td>100 (0)</td>
<td>100 (0)</td>
<td>21 (33)</td>
</tr>
<tr>
<td>Closure dur., ms</td>
<td>60 (72)</td>
<td>45 (10)</td>
<td>111 (17)</td>
</tr>
<tr>
<td>Release dur., ms</td>
<td>26 (9)</td>
<td>49 (14)</td>
<td>20 (16)</td>
</tr>
<tr>
<td>Amp. diff., dB</td>
<td>24 (7)</td>
<td>17 (7)</td>
<td>32 (10)</td>
</tr>
<tr>
<td>Peak (part 1), Hz</td>
<td>2,997 (1,708)</td>
<td>5,715 (1,156)</td>
<td>3,977 (1,674)</td>
</tr>
<tr>
<td>Peak (part 2), Hz</td>
<td>–</td>
<td>4,566 (1,471)</td>
<td>–</td>
</tr>
<tr>
<td>C2 closure, ms</td>
<td>62 (8)</td>
<td>68 (114)</td>
<td>–</td>
</tr>
<tr>
<td>F2 on, Hz</td>
<td>1,888 (204)</td>
<td>2,004 (177)</td>
<td>–</td>
</tr>
</tbody>
</table>
A few speaker-particular differences should be noted. Speaker 3 had considerably greater /t/ versus /tʃ/ peak frequency differences than the other speakers: 5,758 versus 534–2,835 Hz at point 1; these were accompanied by smaller F2 differences prior to the stop closure: 280 versus 395–586 Hz for other speakers. Unlike the other speakers, speaker 4 showed no amplitude differences between /t/ and /tʃ/. Speakers 2 and 3 showed no differences in F2 at the offset of the following /k/. Only three speakers (speakers 2–4) showed a substantial difference between /t/ and /tʃ/ in the duration of the following /k/ (9–13 ms).

To determine to what extent the palatalized stop /tʃ/ different from the 'true' affricate /t∫/, a comparison was made between the two consonants [as produced in ‘Nadja’ and in a filler word ‘savings’ (colloq.)]. /tʃ/ had a shorter release duration \[F(1, 11) = 9.409, p \leq 0.05\], a higher amplitude difference \[F(1, 11) = 19.570, p \leq 0.001\] and a higher peak frequency at point 1 \[F(1, 11) = 43.570, p \leq 0.001\] than /t∫/.

The spectrum of /tʃ/ showed considerable energy at a wider range, between 2,500 and 8,000 Hz, often with two prominent peaks at the lower and upper ends of the range. There were no other significant differences between the two, including no F2 differences.

4.2.2 Context

There was an effect of context (_k, _n, and _s; for /t/) on F2 frequency at the mid-point of the vowel \[F(1, 59) = 3.467, p \leq 0.05\]. A post hoc test revealed that F2 was higher for the /k/ context than for the /n/ context \[p \leq 0.05\]; however, it appeared to have been influenced by preceding palatalized consonants. There were no significant contextual F2 differences at the offset of the vowel \[F(1, 59) = 0.078, p \geq 0.925\]. The effect of the context on % released of /t/ was significant \[F(1, 59) = 273.986, p \leq 0.001\], with significant differences between all three contexts \([p \leq 0.01–0.001\). While /t/ before /k/ was always audibly released, the same consonant was released less commonly before /n/ (21%), and never released before /s/ (at least independently from the fricative noise of /s/). Before /n/, speakers 3, 5, and 6 did not release /t/ at all, while there was some variation in % released among the other speakers (75% for speaker 2 and 25% for speakers 1 and 4).

There was a context effect on closure duration \[F(1, 59) = 86.765, p \leq 0.001\], indicating a longer closure before /n/ than before /k/ and /s/ \([p \leq 0.001\). Comparing /t/ releases before /k/ and /n/ in the production of speaker 2 (who released before /n/ most frequently; see below), release duration was longer before /k/ than before /n/ \[F(1, 7) = 16.148, p \leq 0.01\], while showing no significant differences in amplitude and peak frequency.

4.3 Discussion

The results of the experiment showed that the word-internal contrast between /t/ and /tʃ/ before /k/ was well differentiated acoustically. VC transition differences were very robust, extending throughout at least half of the vowel /a/ (higher F2 for /tʃ/ than for /t/). The average F2 difference was 153 Hz at the midpoint of the vowel /a/ and 123 Hz.

4.2.2 Context

\[
\begin{align*}
\text{zā'natʃ'kā} & \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \qua
466 Hz at its offset (which is somewhat greater than has previously been reported for CV transition differences in the context of /a/ [Fant, 1970; Bondarko 1977]; see section 2.1).

Both /t/ and /t\textsuperscript{jsuperscript} before /k/ were always released [Jones and Ward, 1969; Kuznetsova, 1969]. Release differences between the two were manifested in release duration (longer for /t\textsuperscript{jsuperscript} than for /t/), release amplitude (louder for /t\textsuperscript{jsuperscript} than for /t/), and peak frequency of the burst (higher for /t\textsuperscript{jsuperscript} than for /t/) [see Fant, 1970; Bolla, 1981, on word-initial /t/ and /t\textsuperscript{jsuperscript}]. Differences in release duration were accompanied by the reverse differences in closure duration, reflecting different degrees of affrication of /t/ and /t\textsuperscript{jsuperscript} [Halle, 1959/1971]. While the release of /t/ accounted for 30% of the average overall duration of the stop (closure /H11001 release), the release of /t\textsuperscript{jsuperscript} accounted for 52% of its overall duration; the latter showed considerable interspeaker variation [Kuznetsova, 1969; Bolla, 1981]. This suggests that the relative degree of affrication ([t] vs. [ts\textsuperscript{jsuperscript} or [ts\textsuperscript{jsuperscript}]) could be another important acoustic property distinguishing the contrast, as well as contrasting /t/ with stops at other places of articulation. In fact, the quality and duration of this release appear to be the only properties distinguishing the syllable-final /t\textsuperscript{jsuperscript} and /t\textsuperscript{∫}\textsuperscript{jsuperscript}.

Further, some acoustic information about the contrast was also provided by the closure duration of the following /k/ (longer after /t\textsuperscript{jsuperscript} than after /t/) and F 2 at its offset (higher for /t\textsuperscript{jsuperscript} than for /t/). These differences were relatively small and not always consistent across speakers.

In sum, the contrast between /t/ and /t\textsuperscript{jsuperscript} is signaled by the VC transitions of the preceding vowel, stop releases, as well as by the following consonant. The phonetic realization of the contrast thus spans a number of consonants and vowels. For example, the contrast between the word pairs [katk\textsuperscript{vinverted}] versus [kat\textsuperscript{jsuperscript}k\textsuperscript{vinverted}] and [vz\textsuperscript{jsuperscript}at\textsuperscript{vinverted}] versus [s\textsuperscript{jsuperscript}at\textsuperscript{jsuperscript}k\textsuperscript{vinverted}] can be narrowly transcribed as [katk\textsuperscript{vinverted}] versus [ka\textsuperscript{plussubscript}j k\textsuperscript{plussubscript}{vinverted}] and [vz ji a\textsuperscript{plussubscript}tk\textsuperscript{vinverted}] versus [s ji\textsuperscript{ash}ji k\textsuperscript{plussubscript}{vinverted}]. These contextual effects reflect the articulatory properties of the tongue body raising and fronting gesture (palatalization), which tends to have substantial extent in time, overlapping with the gestures of preceding and following vowels and consonants [Kochetov, in press].

Unlike before /k/, the plain stop /t/ before /n/ was rarely audibly released. When released, its release was shorter in duration compared to the release of /t/ before /k/. Further, /t/ before /s/ never showed a release independent of the fricative noise of the following fricative. Even if the stop were released, its release could have been masked by the high-amplitude noise of /s/.

In sum, the results of the experiment have established the acoustic cues to the word-internal contrast /t/ versus /t\textsuperscript{jsuperscript} before /k/, showing that the contrast is well-cued in this particular context. The next step is to determine which acoustic parameters – releases or VC transitions – are more important for the perceptual identification of the contrast in the context. This question will be addressed in experiment 2, described in the next section. The next step is to determine whether this acoustic finding can be extended to both /t/ and /t\textsuperscript{jsuperscript} across word boundary. This question will be addressed in experiment 3.

5 Experiment 2: Perception of Word-Internal /t/ and /t\textsuperscript{jsuperscript} before /k/
5.1 Method

PsycScope

5.2 Results

= p < = p <
of palatalization were identified as ‘palatalized’ more often than stimuli with the same releases or no releases ($p < 0.001–0.05$). Between the latter two types, stimuli with the same releases were identified as ‘palatalized’ more often than stimuli with no releases ($p < 0.05$). There was, however, a significant VC transitions release interaction $[F(2, 71) = 35.390, p < 0.001]$, indicating that the above-mentioned difference between plain and palatalized VC transitions was not significant for stimuli with different releases; the difference between stimuli with different releases and same/no releases was limited to plain VC transitions; the difference between stimuli with same releases and no releases was limited to palatalized VC transitions. In sum, significant differences involved the following stimuli: [ka(t/corner)tk/vinverted] and [ka(t/corner)tk/corner/vinverted] versus [ka(t/corner)tk/vinverted] and [ka(t/corner)tk/corner/vinverted]; [ka(t/corner)tk/corner/vinverted] versus [ka(t/corner)tk/vinverted] and [ka(t/corner)tk/corner/vinverted], and [ka(t/corner)tk/corner/vinverted] versus [ka(t/corner)tk/corner/vinverted].

With respect to RT, there was a main effect of release $[F(2, 71) = 3.818, p < 0.05]$: it took listeners longer to process stimuli with releases different from VC transitions than those with the same releases as transitions ([ka(t/corner)tk/vinverted] and [ka(t/corner)tk/corner/vinverted] vs. [ka(t/corner)tk/vinverted] and [ka(t/corner)tk/corner/vinverted]); $p < 0.05$). There was also a marginal difference between the former stimuli and those with no releases ([ka(t/corner)tk/vinverted] and [ka(t/corner)tk/corner/vinverted] vs. [ka(t/corner)tk/vinverted] and [ka(t/corner)tk/corner/vinverted]); $p < 0.058$). The factor VC transitions was not significant $[F(1, 71) = 0.891, p < 0.349]$. There were no significant interactions.

Fig. 1. 

Identification rates (0/1/t, 1/t/\) (a) and RT (ms) (b) for 6 stimuli based on the minimal pair [katk/vinverted]-[kat/corner/k/vinverted]: the plain-plain [ka(t/corner)tk/vinverted] ([t/t]), plain-plain/unreleased [ka(t/corner)t/corner/k/vinverted] ([t/t*]), palatalized-plain [ka(t/corner)tk/vinverted] ([t/corner][t]), palatalized-palatalized/unreleased [ka(t/corner)t/corner/corner/k/vinverted] ([t/corner][t/corner]*]), plain-palatalized [ka(t/corner)tk/corner/vinverted] ([t/corner][t/corner]), and palatalized-palatalized [ka(t/corner)t/corner/corner/k/vinverted] ([t/corner][t/corner][t/corner]).
5.3 Discussion

The results showed that a combination of the plain VC transitions and the plain release ([ka(t/corner)tk/vinverted]) gave a consistent percept of the plain /t/; a combination of the palatalized VC transitions and the palatalized release ([ka(t/corner)tk/vinverted]) gave a consistent percept of the palatalized /t/. Removing the plain release from the signal ([ka(t/corner)tk/vinverted]) had no effect on the percept of /t/. However, removing the palatalized release ([ka(t/corner)tk/vinverted]) had a considerable perceptual effect, often leading to a percept of /t/ (over 40% of the time). The latter case was also characterized by a higher RT. This shows that the palatalized release is important to the identification of /t/ and is expected by listeners. In the absence of this release, some information about the secondary articulation of the consonant is still provided by the palatalized VC transitions. The plain release per se is not crucial to the identification of /t/.

A combination of the plain VC transitions and the palatalized release ([ka(t/corner)tk/vinverted]) gave a strong percept of the palatalized /t/ (over 90% of the time). A combination of the palatalized VC transition and the plain release ([ka(t/corner)tk/vinverted]) gave a percept of the plain /t/ much less frequently (only 40% of the time). Both cases were characterized by a higher RT. This suggests that the palatalized release is somewhat more salient than the palatalized transitions (when juxtaposed to the plain transitions and releases, respectively). Both palatalized transitions and releases are more salient than plain transitions and releases. Any combination of conflicting cues is difficult for listeners to integrate.

In sum, the results of the experiment confirmed the prediction that acoustic information available in the word-internal context a_k (as established in experiment 1) renders the contrast /t/ versus /t/ reliably recoverable. The results also revealed that releases were particularly important for the identification of /t/, while VC transitions were important for the identification of both /t/ and /t/.

Having determined that, we are now in a position to directly test the hypothesis of licensing by cue by examining acoustic and perceptual properties of the stops across word boundaries.

6 Experiment 3: Production of Word-Final /t/ and /t/ before Word-Initial /k/, /n/, and /s/

6.1 Method

Subjects employed in the experiment were the same 6 native speakers of Russian as in experiment 1. Test words presented to the speakers consisted of (near-)minimal word pairs containing the phonemes /t/ and /t/ word finally followed by words beginning with the consonants /k/, /n/, and /s/ (table 4). For all phrases, immediate vowel contexts and stress were held constant: the vowel /a/ preceding the target consonants was stressed; the following vowel /a/ was unstressed.

The data collection was the same as in experiment 1. The total number of recorded target words obtained was 1,944 (12 words/3 contexts/3 repetitions/6 speakers). Only the second production of each utterance was analyzed, giving 108 tokens per subject (a total of 648 tokens). Acoustic

Phonetica 2006 – 127
Table 4. Test words used in experiment 3

<table>
<thead>
<tr>
<th>a_k</th>
<th>[ʼmatʰkan’duktʰa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>_#</td>
<td>‘conductor’s foul language’</td>
</tr>
<tr>
<td>_#</td>
<td>‘taken off for lunch’</td>
</tr>
<tr>
<td>_#</td>
<td>‘taken sometimes’</td>
</tr>
<tr>
<td>_#</td>
<td>‘cashier’s relative (son-in-law’s father)’</td>
</tr>
<tr>
<td>_#</td>
<td>‘wrinkled some time ago’</td>
</tr>
<tr>
<td>_#</td>
<td>‘a basket of mushrooms’</td>
</tr>
<tr>
<td>_#</td>
<td>‘boss’s foul language’</td>
</tr>
<tr>
<td>_#</td>
<td>‘has been) photographed’</td>
</tr>
<tr>
<td>_#</td>
<td>‘taken for long time’</td>
</tr>
<tr>
<td>_#</td>
<td>‘boss’s relative (son-in-law’s father)’</td>
</tr>
<tr>
<td>_#</td>
<td>‘wrinkled attire’</td>
</tr>
<tr>
<td>_#</td>
<td>‘colleague’s foul language’</td>
</tr>
<tr>
<td>_#</td>
<td>‘taken off the stand’</td>
</tr>
<tr>
<td>_#</td>
<td>‘brother is going to’</td>
</tr>
<tr>
<td>_#</td>
<td>‘combat engineer’s relative (son-in-law’s father)’</td>
</tr>
<tr>
<td>_#</td>
<td>‘wrinkled entirely’</td>
</tr>
<tr>
<td>_#</td>
<td>‘picked up mushrooms’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>a_n</th>
<th>[ʼmatʰkan’duktʰa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>_#</td>
<td>‘conductor’s mother’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to take off for lunch’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to take sometimes’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to call a cashier’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to wrinkle some time ago’</td>
</tr>
<tr>
<td>_#</td>
<td>‘a basket again’</td>
</tr>
<tr>
<td>_#</td>
<td>‘boss’s mother’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to photograph’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to take for long time’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to call the boss’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to wrinkle attire’</td>
</tr>
<tr>
<td>_#</td>
<td>‘picked up again’</td>
</tr>
<tr>
<td>_#</td>
<td>‘colleague’s mother’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to take off the stand’</td>
</tr>
<tr>
<td>_#</td>
<td>‘going to take’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to call a combat engineer’</td>
</tr>
<tr>
<td>_#</td>
<td>‘to wrinkle entirely’</td>
</tr>
<tr>
<td>_#</td>
<td>‘picked up again’</td>
</tr>
</tbody>
</table>

released, stop closure and release duration, amplitude difference

following consonant closure duration (C2 closure)

following vowel F2 frequency

F2 on

preceeding vowel F2 frequency, %

stop release peak frequency
6.2 Results

Table 5 presents the results, means and standard deviations for \( t \) and \( t' \) in three word-boundary consonant contexts: before /k/, /n/, and /s/.

6.2.1 Consonant and Context

There was a significant effect of consonant on F2 both at the midpoint of the vowel \( [F(1, 207) = 26.918, p = 0.001] \) and at its offset \( [F(1, 209) = 262.034, p = 0.001] \): F2 values were higher before /t/ than before /t'. The factor context was not significant at either point \( [F(2, 207) = 0.852, p = 0.428; F(2, 209) = 0.675, p = 0.510] \). There were effects of consonant and context on closure duration \( [F(1, 167) = 21.378, p = 0.001; F(2, 167) = 15.206, p = 0.001] \): the closure was longer for /t/ than for /t', as well as longer before /n/ and /s/ than before /k/ \( (p = 0.001–0.01) \). There was, however, a consonant-context interaction \( [F(2, 167) = 4.385, p = 0.05] \), indicating that the consonant difference was not significant before /s/.

The measure of % released was not affected by the factor consonant \( [F(1, 205) = 1.560, p = 0.213] \), while it was affected by the factor context \( [F(2, 205) = 131.729, p = 0.001] \). Both stops were almost always released before /k/ (almost 100%; 1 unreleased token of /t'/); they were considerably less frequently released before /n/ (38%), and rarely released before /s/ \( (p = 0.001) \). Before /n/, speaker 6 did not release the stops at all, while there was considerable variation in % released among the other speakers: 75, 64, 58, 10, and 17% for speakers 1–5, respectively. Before /s/, only speaker 5 released the stops (55% of the time). There were effects of consonant on release duration and amplitude difference \( [F(1, 109) = 5.774, p = 0.05; F(1, 109) = 9.396, p = 0.01] \): it was longer and smaller \( (/louder) \) for /t/ than for /t'. (These differences were found for all speakers, except speaker 4.)

The factor consonant did not significantly affect peak frequency at point 1 \( [F(1, 107) = 0.213, p = 0.645] \), although the mean peak frequency was somewhat higher.

---

<table>
<thead>
<tr>
<th></th>
<th>a_#k</th>
<th>a_#n</th>
<th>a_#s</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2 mid, Hz</td>
<td>1,533 (244)</td>
<td>1,724 (247)</td>
<td>1,608 (226)</td>
</tr>
<tr>
<td>F2 off, Hz</td>
<td>1,424 (198)</td>
<td>1,895 (207)</td>
<td>1,455 (216)</td>
</tr>
<tr>
<td>% released</td>
<td>100 (0)</td>
<td>97 (16)</td>
<td>27 (45)</td>
</tr>
<tr>
<td>Closure dur., ms</td>
<td>57 (13)</td>
<td>36 (15)</td>
<td>76 (23)</td>
</tr>
<tr>
<td>Release dur., ms</td>
<td>24 (9)</td>
<td>42 (18)</td>
<td>26 (14)</td>
</tr>
<tr>
<td>Amp. diff., dB</td>
<td>26 (8)</td>
<td>18 (9)</td>
<td>36 (2)</td>
</tr>
<tr>
<td>Peak (part 1), Hz</td>
<td>4,599</td>
<td>5,721</td>
<td>5,669</td>
</tr>
<tr>
<td>(2,040) (1,151)</td>
<td>(2,075) (1,624)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak (part 2), Hz</td>
<td>–</td>
<td>4,903 (1,646)</td>
<td>–</td>
</tr>
<tr>
<td>C2 closure, ms</td>
<td>54 (15)</td>
<td>57 (17)</td>
<td>39 (17)</td>
</tr>
<tr>
<td>F2 on, Hz</td>
<td>1,568 (147)</td>
<td>1,638 (149)</td>
<td>1,424 (219)</td>
</tr>
</tbody>
</table>
for \(/t^{superscript}/\) than for \(/t/\) (by over 1,000 Hz). The two also differed in the overall distribution of energy: relatively compact \(/t^{superscript}/\) versus relatively diffuse \(/t/\). Measurements of the \(/t/\) release were not possible at point 2. The effect of context on releases was considered for speakers who showed a higher than 50% release rate before /n/ or /s/ (speakers 1, 2, 3 and 5). Compared to releases before /k/, releases before /n/ had a higher amplitude difference (\(H11005\) less loud); releases before /s/ had a shorter duration and a higher amplitude difference. There were no other contextual differences.

The duration of the following consonant was not affected by the factor consonant \(F(1, 204) = 0.041, p = 0.839\], while \(F2\) at the offset of the following consonant was \(F(1, 208) = 21.581, p = 0.001\]. \(F2\) was higher after \(/t^{superscript}/\) than after \(/t/\) (for all speakers, except speaker 3).

6.2.2 Boundary

The factor boundary (\(a_#k\) vs. \(a_k\)) did not affect \(F2\) frequency at either of the two compared points – the vowel midpoint and its offset \(F(1, 145) = 2.047, p = 0.155\); \(F(1, 146) = 0.065, p = 0.799\]. There was an effect of boundary on closure duration \(F(1, 145) = 11.980, p = 0.01\]: word-final consonants had a shorter closure duration. They also had a marginally shorter release duration \(F(1, 145) = 3.777, p = 0.054\]. Release amplitude was not affected \(F(1, 145) = 0.912, p = 0.341\]. There was a boundary effect on peak frequency at point 1 \(F(1, 145) = 9.589, p = 0.01\], indicating that word-final bursts had a higher peak frequency than word-internal bursts. This difference, however, was limited to \(/t/\) \(boundary\) \(consonant\) interaction: \(F(1, 145) = 9.589, p = 0.01\].9 There was no boundary effect on peak frequency at point 2 \(\text{only;}\ F(1, 145) = 0.630, p = 0.429\].

Compared to the word-internal \(/t/\) before /n/, the word-final \(/t/\) before /n/ had a shorter closure duration \(F(1, 21) = 17.019, p = 0.01\] and a somewhat higher \(F2\) frequency at the midpoint of the preceding vowel \(F(1, 44) = 14.837, p = 0.001\]. No other significant boundary differences were found \(F2\ off: F(1, 44) = 0.447, p = 0.507\); release duration: \(F(1, 15) = 0.517, p = 0.484\); amplitude difference: \(F(1, 15) = 1.565, p = 0.231\]; peak frequency, point 1: \(F(1, 14) = 2.543, p = 0.135\]. There were no significant differences between the word-internal \(/t/\) before /s/ and the word-final \(/t/\) before /s/ \(F2\ mid: F(1, 42) = 1.833, p = 0.183\); \(F2\ off: F(1, 42) = 0.040, p = 0.843\); closure duration: \(F(1, 42) = 0.449, p = 0.507\].

6.3 Discussion

The results of the experiment showed that the word-final contrast \(/t/\) versus \(/t^{superscript}/\) was differentiated before the word-initial /k/ by VC transitions (\(F2\) differences throughout at least half of the vowel /a/), releases (release duration, amplitude, and peak frequency of the burst), as well as by \(F2\) at the onset of the following vowel /a/. An important finding was that both stops were released almost 100% of the time \(see\ Zsiga, 2000,\ on\ the\ Russian\ /d/#/k/\ sequence\). Thus, in terms of the overall acoustic information available about the contrast, this word-final context was very similar to the word-internal context before /k/ (\(a_k\); see section 4.2.1), with minor differences in duration and peak frequency.

9 The boundary difference for \(/t/\) was in part due to several tokens from speaker 3 who showed relatively weak word-final /t/ releases with peaks at high frequencies (6,000–9,000 Hz).
The two stops before /n/ and /s/ were differentiated by VC transitions of the magnitude comparable to the context before /k/, as well as by minor formant differences after the following consonant. However, unlike the stops before /k/, the stops before /n/ were more often unreleased than released, and the stops before /s/ were rarely released [see Zsiga, 2000, on Russian /d/#/t/ sequence]. When released, the stops before /n/ and /s/ were not differentiated by release duration or amplitude to the same extent as releases before /k/.

In sum, the word-boundary contexts before /k/, /n/, and /s/ differ at least in one important respect: while the hetero-organic context before /k/ favors saliently released stops, the two homorganic contexts do not. This finding confirms our tentative conclusions about differences among word-internal contexts (section 4.3). Having determined that the contexts before /k/, /n/, and /s/ differ acoustically, we still need to establish whether these acoustic differences are relevant to the identification of the contrast.

7 Experiment 4: Perception of Word-Final /t/ and /t/ before /k/, /n/, and /s/

The goal of this experiment was to determine whether the perception of the contrast between the word-final /t/ and /t/ (and particularly the perception of /t/) is affected by the context, the following word-initial /k/, /n/, and /s/.

7.1 Materials, Procedure, and Analysis

The materials included the minimal pair [mat]–[mat/]/ used in utterances with words beginning with /k/, /n/, and /s/, as shown in table 6. These utterances were a subset of the stimuli from experiment 3 pronounced by speaker 1. Acoustic analysis showed that this subject released /t/ and /t/ 100% of the time before /k/, 67 and 83% of the time, respectively, before /n/, and never before /s/. In the tokens of [mat] and [mat/] before /n/ used for the current study, the stops were released (see Appendix).

Additional stimuli were created by copying the portion of the signal that included the nasal murmur of /m/ and the whole vowel /a/ (from its onset to the offset) of the word [mat] and substituting it for the corresponding portion of the item [mat/] (for all 3 utterance pairs in table 6). This resulted in stimuli with conflicting acoustic cues: ‘plain’ VC transition and ‘palatalized’ release (if present; [ma(t#k)n#tʃ#n#k], [ma(t#n)tʃ#al#n#k], and [ma(t#s)trud#n#k]). This set of stimuli was created to determine the net contribution of the release to signaling the contrast, particularly in the acoustically less informative contexts (_#n and _#s).
Participants, stimuli presentation, and data collection were the same as in experiment 2. There were a total of 216 responses (9 stimuli \( /t/ \) repetitions 2 repetitions /\( /k/ \) \( /n/ \) \( /s/ \) 12 listeners), or 18 responses per listener. As before, the analysis was based on mean identification rates (ratios of \( 't/ \) responses) and RT per subject. (Twenty-two responses were excluded from the analysis based on the criteria described for experiment 2.) The repeated-measures ANOVA involved two within-subject factors: consonant (/t/, /t/ \( /k/ \), and \( /t/ \) \( /n/ \) \( /s/ \)) and context (_#k_, _#n_, and _#s_). The measures were identification rate and RT.

### 7.2 Results

Mean identification rates for the three consonant types (/t/, /t/ \( /k/ \), and the edited /t/ \( /n/ \) \( /s/ \), \( /t/ \) \( /n/ \) \( /s/ \)) are plotted in figure 2a). As in experiment 2, the value of 1 indicates that all responses to a given stimulus were \( 't/ \); the value of 0 indicates that all responses to a given stimulus were \( 't/ \). Mean RT for the three consonant types in the three contexts are given together in figure 2b.

The results showed that, with respect to the identification rate, there were main effects of consonant \( F(1, 11) = 231.738, p < 0.001 \) and context \( F(1, 11) = 6.760, p < 0.01 \). Post hoc tests revealed that identification rates for all consonant types (/t/, /t/ \( /k/ \), and the edited /t/ \( /n/ \) \( /s/ \), \( /t/ \) \( /n/ \) \( /s/ \)) were significantly different from each other \( p < 0.01–0.001 \). There was a significant consonant context interaction \( F(2, 22) = 27.308, p < 0.001 \), indicating that differences between the contexts were significant for the edited /t/ \( /n/ \) \( /s/ \) (higher 't/ identification rates before /k/ than before /n/ and /s/; significant contextual differences were limited to the difference between the /k/ and /n/ contexts \( p < 0.01 \)).

With respect to RT, there was no main effect of consonant \( F(1, 11) = 0.544, p > 0.05 \). There was a main effect of context \( F(1, 11) = 4.772, p < 0.05 \) and a significant consonant context interaction; \( F(2, 22) = 5.359, p < 0.01 \). The interaction was due to the difference between the edited /t/ \( /k/ \) (higher RT) and /t/ \( /n/ \) \( /s/ \) (lower RT) in the context of /k/, as well as the difference between the unedited /t/ \( /n/ \) \( /s/ \) before /k/ and /s/ (higher RT) and the same consonant before /s/ (lower RT).

### 7.3 Discussion

The results of the experiment showed that, regardless of the context, correct identification of /t/ and /t/ \( /k/ \) was high. It was near-perfect in the contexts before /k/ and /n/; it was 100% for /t/ \( /k/ \) and 75% for /t/ \( /n/ \) in the context before /s/). The latter effect can be attributed to the overall higher F2 values during the VC transition in the stimulus \( /t/ \) \( /k/ \), compared to the other stimuli (see Appendix). The contexts differed strikingly in the identification of the edited, plain-palatalized \( /t/ \) \( /k/ \). While \( /t/ \) \( /k/ \) \( /n/ \) \( /s/ \) was identified as 'palatalized' most of the time before /k/ (showing a high RT), it was identified as 'plain' in the absolute majority of responses before /n/ and /s/. This result was expected: in the absence of palatalized transitions, the information about /t/ \( /k/ \) was still provided by the salient (relatively long and loud) release (see Appendix). The release of /t/ \( /n/ \) \( /s/ \) before /k/ was present in the signal and classified as 'audible', yet it was obviously less salient – shorter and less loud than the release before /k/. This apparently was sufficient to render the release before /n/ imperceptible under the current experimental conditions. Other acoustic differences, such as F2 differences at the offset of /n/, did not affect the results. Before /s/, /t/ \( /s/ \) in the stimulus was not released.
Testing Licensing by Cue

Identification rate

Context: a_#k

Context: a_#n

Context: a_#s

Fig. 2. Identification rates (0/1005/t/, 1/1005/t/ with the plain transition ([t][t] before /k/, /n/, and /s/, in the minimal pair [mat]-[mat^]{superscript}.)
8 Experiment 5: Production of Utterance Final /t/ and /t’/

8.1 Method

8’z, it’1

C1 closure and release duration
Preceding vowel duration

% released,

amplitude difference

Preceding vowel F2 frequency

Stop release peak frequency
The collected data were analyzed in an ANOVA with the independent variables consonant (/t/ and /t/) and context (a_#, u_#, i_#). The dependent variables were the acoustic measurements mentioned above. To compare the results to those obtained in experiments 1 and 3, an additional ANOVA was run with the variables boundary (a_##, a_k, and a_#k) and consonant. (When comparing across contexts, words with comparable preceding consonants were selected, whenever possible).

8.2 Results

Table 8 presents the results, means and standard deviations for the utterance-final /t/ and /t/jsuperscript/ in three vocalic contexts: after /a/, /u/, and /i/.

8.2.1 Consonant and Context

The factor consonant did not significantly affect F 2 at the midpoint of the vowel \[F(1, 72) = 3.453, p = 0.068\], while showing a nonsignificant trend towards higher values for /t/jsuperscript/ than for /t/. The two consonants differed in the context of /a/ by 137 Hz \[F(1, 23) = 5.746, p = 0.05\]; the difference was relatively small in the contexts of /u/ and /i/ (82 and 25 Hz, respectively; 135 Phonetica 2006 63:113 – 148 Testing Licensing by Cue Table 8.}

<table>
<thead>
<tr>
<th>a_#</th>
<th>[mat] 'foul language'</th>
<th>[mat'] 'mother'</th>
</tr>
</thead>
<tbody>
<tr>
<td>z_i</td>
<td>t 'married'</td>
<td>t' 'to drive'</td>
</tr>
<tr>
<td>_#</td>
<td>t 'fetters'</td>
<td>t' 'way'</td>
</tr>
<tr>
<td>_#</td>
<td>t 'blown'</td>
<td>t' 'to blow'</td>
</tr>
<tr>
<td>_#</td>
<td>t 'beaten'</td>
<td>t' 'to beat'</td>
</tr>
<tr>
<td></td>
<td>t 'imagines (onself)'</td>
<td>t' 'thread'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V duration, ms</td>
<td>139 (23)</td>
<td>146 (25)</td>
</tr>
</tbody>
</table>
8.2.2 Boundary

\[ p < \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]

\[ = \]

\[ p < \]
8.3 Discussion

The results of the experiment showed that the utterance-final contrast /t/ versus /t/ was differentiated by both VC transitions and consistent releases. Differences in VC transitions varied substantially as a function of the vowel context. In the context of /a/, F2 differences extended for almost half of the vowel (i.e. on average for about 70 ms). In the context of /u/, F2 differences were limited to the last two points, thus extending for a little more than one fourth of the vowel (i.e. on average for 30 ms). In the context of /i/, F2 differences were found primarily at the offset of the vowel. In the /a/ and /u/ contexts, F2 transitions were more rising for /t/ than for /t/; in the /i/ context, they were level (or slightly falling) for /t/ and falling for /t/.

F2 values at the offset of /a/ were very close to the locus ('terminal') values predicted for plain and palatalized denti-alveolars [Fant, 1970; see section 2.1]; F2 values at the offset of /u/ were considerably lower, and F2 values at the offset of /i/ were considerably higher than the locus values. Assuming that the rising F2 during the VC transition is one of the main cues to syllable-final palatalized consonants [Halle 1959/1971, pp. 151–153],10 our findings suggest that the low central vowel /a/ provides more acoustic information about the palatalized stop /t/ (and therefore about its contrast with /t/) compared to the high back vowel /u/; the latter provides more acoustic information compared to the high front vowel /i/. It is true, however, that the contrast may still be adequately distinguished by differences in transitions to both consonants (negative transitions vs. level transitions).

Further, both stops had somewhat longer and louder releases after the inherently shorter high vowels /u/ and /i/. In fact, /t/ releases after /u/ and /i/ were on average louder than the respective vowels. The longer duration and higher amplitude of the utterance-final stops (compared to the same stops in the other two contexts) are apparent effects of utterance-final gestural 'strengthening' [Fougeron and Keating, 1997], which can be regarded as phonetic enhancement of the contrast in a prominent prosodic position. Note that considerable periods of frication in releases were accompanied by voiceless [ibarred]- or [i]-like off-glides [Fant, 1970; Bondarko, 1981], and therefore, can be narrowly transcribed as [t/ibarred/underring] and [tsji/underring] (e.g. [mati/underring] vs. [ma/plussubscript/itsji/underring], [dut/ibarred/underring] vs. [duitsji/underring], and [bji/ibarred/underring] vs. [bjitsji/underring]). Such a realization of releases in the utterance-final context makes them phonetically comparable to short voiceless vowels [ibarred/underring] and [i/underring], quite plausibly, phonetic traces of the historically deleted following high vowels [Carlton, 1990]. The apparent lack of robust differences in peak frequency during releases of /t/ and /t/ (at points 2 and 3) may reflect considerable
9 Experiment 6: Perception of Utterance Final /t/ and /t/ after /a/, /u/, /i/

9.1 Method

The goal of this experiment was twofold: (a) to determine the relative importance of VC transitions and stop releases in the perception of the utterance-final /t/ and /t/ and (b) to determine whether the perception of the contrast is affected by vowel quality.

9.1.1 Method

Test words for the first part of the experiment consisted of the minimal pair [mat]-[mat], as pronounced by speaker 1 in the production of experiment 5 (see Appendix). As in experiment 2, the items were excised from carrier phrases and segmented into two components. The first component in each case included the nasal murmur of the initial /m/ and the entire following vowel /a/ (from its onset to the offset), that is the sequences [ma[t/corner]) and [ma(t/corner]) (where the stops in parentheses indicate VC transitions). The second component included the closure and release of the target stop, that is [t] and [t/corner].

The four components were spliced together in four possible combinations: (a) the plain-plain [ma(t/corner)t], (b) the palatalized-plain [ma(t/corner)t], (c) the plain-palatalized [ma(t/corner)t], and (d) the palatalized-palatalized [ma(t/corner)t]. As in experiment 2, the stimuli included two tokens with removed releases: (e) the plain-plain/unreleased [ma(t/corner)t/corner] and (f) the palatalized-palatalized/unreleased [ma(t/corner)t/corner]. This set of stimuli was designed to determine the relative importance of VC transitions (or preceding vowels in general in [ma(t/corner]) and [ma(t/corner)]) and releases.

The second part of the experiment involved three minimal pairs – [mat]-[mat], [dut]-[dut], and [b/corner]-[b/corner], as pronounced by speaker 1 in the production of experiment 5 (see Appendix). Additional stimuli included the forms with removed palatalized releases, that is [mat/corner], [dut/corner] and [bit/corner]. (The stimuli [mat], [mat], and [mat/corner] were used in both part 1 and part 2 of the experiment.) This was done to determine the difference between vowel contexts in the identification of the plain/palatalized contrast.

Participants, stimuli presentation, and data collection were the same as in experiments 2 and 4. As before, the clock for measuring RT was started at the beginning of the stimulus (to include listeners' responses based solely on the VC transition information). For part 1, there were a total of 144 responses (6 stimuli x 2 repetitions x 12 listeners), or 12 responses per listener. For part 2, there were a total of 216 responses (9 stimuli x 2 repetitions x 12 listeners), or 12 responses per listener. The statistical analysis was based on mean identification rates (the ratio of 't/corner' responses) and RT per subject, done separately for part 1 and part 2. The criteria for selecting RT responses were the same as in previous experiments. (In total, 5 responses were excluded from the analysis in part 1 and 13 responses were excluded from the analysis in part 2.) For part 1, an ANOV A was conducted with two independent variables, VC transition (plain and palatalized) and release (same value as transitions, no release, different value from transitions). For part 2, a repeated-measures ANOV A was conducted with the variables consonant (/t/, /t/, and the unreleased /t/) and context (a#, u#, and i#). In both analyses, the measures were identification rate (ratio 't/corner') and RT.
9.2 Results: Part 1

The value of 1 indicates that all responses to a given stimulus were ‘t<sup>3</sup>’; the value of 0 indicates that all responses to a given stimulus were ‘t<sup>2</sup>’. The stimuli in both graphs are ordered from the lowest identification value for ‘t<sup>2</sup>’ to the highest.

With respect to the identification rate, there were main effects of VC transition \[F(1, 71) = 53.975, p < 0.001\] and release \[F(2, 71) = 3.588, p < 0.05\]. Stimuli with palatalized VC transitions were identified by listeners as ‘palatalized’ more often than stimuli with plain transitions. Stimuli with releases having the same palatalization values as VC transitions were identified as ‘palatalized’ more often than stimuli with no releases \[p < 0.05\]; they were not significantly different from stimuli with different transitions \(p < 0.081\). There was, however, a significant VC transition/release interaction \[F(2, 71) = 18.728, p < 0.001\], indicating that the above-mentioned difference between plain and palatalized VC transitions was not significant for stimuli with different releases; the difference between stimuli with the same or no releases was limited to stimuli with palatalized VC transitions. Differences among stimuli with plain

---

**Fig. 3.**

![Identification rates and RT for 6 stimuli based on the minimal pair: plain-plain, plain-plain/unreleased, palatalized-plain, plain-palatalized, palatalized-palatalized/unreleased, and palatalized-palatalized.](image)

---

Phonetica 2006 – 139
9.3 Results: Part 2

\[ p < \quad = \quad p < \]

\[ p < \quad = \quad p < \]

\[ \times \quad = \quad p < \]

\[ \times \quad = \quad p < \]

9.4 Discussion

Fig. 4.
They show that the palatalized release is important to the identification of /t\(^\text{\textsuperscript{\textit{p}}}\) and is expected by listeners. In the absence of this release, the secondary articulation of the consonant is still partly cued by the palatalized VC transitions. The plain release is not crucial to the identification of /t/.

A combination of the plain VC transitions and the palatalized release ([ma(t\(^\text{\textsuperscript{\textit{p}}}\)t\(^\text{\textsuperscript{\textit{p}}}\)]) gave conflicting results, however, with a stronger percept of the plain /t/ (almost 70% of the time). A combination of the palatalized VC transition and the plain release ([ma(t\(^\text{\textsuperscript{\textit{p}}}\)]) gave very similar results: a stronger percept of the plain /t/ (70% of the time). Both cases were characterized by a higher RT. This suggests that the palatalized release in this context is about as salient as the palatalized transitions (when juxtaposed to the plain transitions and releases, respectively). Both palatalized transitions and releases appear to be less salient than the plain transitions and releases. Any combination of conflicting cues is similarly difficult for listeners to integrate. These results are rather different from the word-internal contrast (experiment 2), where the palatalized release was more salient than the palatalized transitions, and both were more salient than the plain transitions and release.

The results of the second part of the experiment showed that, while the vowel context has no apparent effect on the identification of released palatalized stops /t/ and /t\(^\text{\textsuperscript{\textit{p}}}\), it had a considerable effect on the unreleased [t\(^\text{\textsuperscript{\textit{p}}}\)]. Importantly, this effect differed by the vowel context: the unreleased [t\(^\text{\textsuperscript{\textit{p}}}\)] after the back vowels /a/ and /u/ was identified randomly either as 't' or 't\(^\text{\textsuperscript{\textit{p}}}\)'; after the high front vowel /i/, it was consistently misidentified as 't'. The latter effect was expected, however, its magnitude was somewhat surprising.

In terms of RT, the unreleased [t\(^\text{\textsuperscript{\textit{p}}}\)] showed a considerably higher RT. The higher RT for the /a/ context (for /t/ and /t\(^\text{\textsuperscript{\textit{p}}}\)) can be attributed to the longer duration of /a/ compared to /i/ and /u/. In sum, the results confirmed that the palatalized release is crucial to the identification of /t\(^\text{\textsuperscript{\textit{p}}}\), yet it is more important for the context after /i/ than after /a/ and /u/. The latter vowel contexts provide sufficient information about the contrast through VC transition differences, while the VC transition differences in the /i/ context are rather small. This result confirms the observation by Jones and Ward [1969] that VC transitions to palatalized stops in the context of preceding front vowels are barely perceptible [see also Thomson, 1931].

In sum, the results of the perceptual experiment have confirmed that the contrast /t/ versus /t\(^\text{\textsuperscript{\textit{p}}}\) utterance finally (a_#) differs perceptually from the same contrast word-finally before /k/ (a_#k), which gives somewhat different acoustic properties distinguishing the contrast in these contexts. Specifically, VC transitions are more important than releases in the a_# context compared to the a_#k context. Plain releases are also more important in the former context than in the latter. Further, the results showed that acoustic differences between the vowel contexts are important perceptually: the greater amount of information about the contrast provided by the preceding vowels /a/ and /u/ compared to the vowel /i/ (which is to some degree subject to individual variation) results in better recoverability of the contrast in the former contexts compared to the latter context. However, substantial acoustic differences between the /a/ and /u/ contexts did not turn out to have a perceptual effect. The findings of the experiment can be extended to word-internal contexts with different vowel contexts (e.g. a_#k, u_#k, i_#k).

11 The result may be in part attributed to the relatively small F2 differences between /t/ and /t\(^\text{\textsuperscript{\textit{p}}}\) at point 3 in the stimuli (116 Hz; table 13 in the Appendix), as well as to the same direction of the transitions to /t/ and /t\(^\text{\textsuperscript{\textit{p}}}\) (both falling).

It should be noted, however, that the difference in the context of /u/ at point 3 in the stimuli was even smaller and also characterized by transitions similar to each other in direction (both rising).
10 General Discussion

The key prediction of the hypothesis of licensing by cue was that the Russian contrast /t/ versus /t_/ is licensed in some contexts (e.g. _# and _k) because these contexts provide enough information about the contrast (more and better cues); it is neutralized in other contexts (e.g. _n and _s) because these contexts cannot potentially provide the same amount of information about the contrast (fewer and poorer cues or no cues).

Since licensing and neutralization of the contrast in Russian and other Slavic languages make reference to the following consonants or the absence of such segments (e.g. _k and _# vs. _n and _s), the prediction was that the cues at the right edge of the stops (releases) are crucial to the identification of the contrast. Since licensing and neutralization of the contrast did not make any reference to the preceding vowel (e.g. a_, u_, or i_), the prediction was that the cues at the left edge of the stops (VC transitions) are considerably less important for the identification of the contrast (if important at all).

The series of experiments undertaken in this study investigated acoustic and perceptual properties of some of these and of similar word boundary contexts. The results of the experiments were further extended to other relevant contexts, not investigated directly. The results are summarized in table 9 for each context, referring to preceding vowels or following consonants (if present). Note that the word-final context before consonants and the word-internal contexts are grouped together, since their phonetic similarity was established in the study. Each cell briefly characterizes the key cues to the /t/ versus /t_/ contrast in terms of relative acoustic/perceptual qualities of VC transitions and releases. For the sake of comparison, the amount of VC transition information is roughly categorized as 'salient', 'less salient' and 'least salient', based on the corresponding F2 differences of preceding vowels. The amount of release information is categorized as 'salient', 'less salient' and 'no release', based on the corresponding differences in % released, release amplitude, peak frequency, and duration (relative to closure duration). Other cues that do not seem to clearly differentiate the contexts (e.g. following consonants) are omitted. Caution should be taken when considering cells with the contexts not directly investigated in the study (preconsonantal contexts after /u/ and /i/).

Recall that, overall, both VC transitions and releases were found to be important to the identification of the contrast. There was some contextual variation: palatalized VC transitions were more important than palatalized releases in certain utterance-final contexts (e.g. a_# and u_#); they were somewhat less important than palatalized releases in some preconsonantal contexts (e.g. a_k, a_#k, a_#n and a_#s); finally, they were least important in the utterance-final context after /i/. Similarly, plain VC transitions and releases varied with context. Thus, no general statement about the relative salience of the two cues (releases /t/ VC transitions or VC transitions /t/) emerges from the results of the experiments.

An important observation emerges that the contexts in table 9 present a continuum in terms of the amount of information provided about the contrast. The contexts a_# and a_#k (and therefore a_k) in the upper right corner of the table can provide the contrast with 'salient' transitions and 'more or less salient' releases, while the contexts i_#n and i_#s (and therefore i_n and i_s) in the lower right corner of the table can provide only 'least salient' transitions. Other contexts are in between these two extremes. How does the relative amount of acoustic information correspond to the sites of neutralization of the contrast? The contexts where the /t/ versus /t_/ contrast is maintained and neutralized within words in Russian are repeated in table 10.
Comparing tables 9 and 10, there is a very good correspondence between the amount of information provided by releases (the horizontal dimension in table 9) and the contexts where the contrast is maintained or neutralized. The contexts where the two stops can be relatively saliently released are in fact the contexts where the Russian /t/ versus /t'/ contrast is maintained within words. The contexts where the stops' releases are substantially less salient or absent altogether are the contexts where the language neutralizes the contrast within words. Further, the relative salience of utterance-final (_#) releases compared to releases before hetero-organic consonants (_k and _#k) corresponds to the other common distributional pattern of neutralization found in other Slavic languages: the maintenance of the contrast word-finally and its neutralization before consonants, regardless of place of articulation.

At the same time, there appears to be absolutely no correspondence between the amount of information provided by VC transitions (the vertical dimension in table 9) and the contexts where the contrast is maintained or neutralized in Russian. (Recall that values for /u/ and /i/ VC transitions were experimentally established for the utterance-final context only and are hypothesized to be similar in the other contexts.) Thus, the contrast is maintained (_# and _k) regardless of whether VC transitions are 'salient' or not; similarly, the contrast is neutralized (_n and _s) regardless of the quality of the transitions. Neither in Russian (the standard or other dialects) nor in other Slavic languages.

<table>
<thead>
<tr>
<th>_#</th>
<th><em># ~</em></th>
<th><em># ~</em></th>
<th><em># ~</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>a_</td>
<td>Salient transitions</td>
<td>Salient transitions</td>
<td>Salient transitions</td>
</tr>
<tr>
<td></td>
<td>more information</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>_#</th>
<th><em># ~</em></th>
<th><em># ~</em></th>
<th><em># ~</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>less information</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 9.**

<table>
<thead>
<tr>
<th>_#</th>
<th><em># ~</em></th>
<th><em># ~</em></th>
<th><em># ~</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>a_</td>
<td>/t/ vs. /t'/</td>
<td>/t/ vs. /t'/</td>
<td>only /t/</td>
</tr>
<tr>
<td></td>
<td>contrast maintained</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>_#</td>
<td>_#</td>
<td></td>
</tr>
<tr>
<td></td>
<td>contrast maintained</td>
<td>contrast neutralized</td>
<td></td>
</tr>
</tbody>
</table>

**Table 10.**
do we seem to find the patterns of neutralization referring to VC transitions. For exam-
ple, we do not find patterns where the contrast is maintained after /a/ but neutralized after
/i/, either word-finally, or before hetero-organic or homorganic consonants (e.g.
a_k [H11022] i_k or a_n [H11022] i_n).

While the clear correspondence between the salience of releases and neutraliza-
tion patterns provides strong support for the hypothesis of licensing by cue, the lack of
any correspondence between the salience of VC transitions and neutralization patterns
is somewhat puzzling. VC transitions are almost as informative, or, in some contexts,
even more informative than releases. Unlike palatalized stop releases, palatalized VC
transitions are the only acoustic property shared by all syllable-final palatalized conso-
lants [Fant, 1970], and referring to the same articulatory gesture of palatalization.
Unlike stop releases, transitions are known to be considerably less susceptible to back-
ground noise, and thus potentially more informative under natural speech conditions
[see Wright, 2001, for discussion]. Yet, these factors do not seem to correlate with the
roles VC transitions and releases play in phonology. It appears that, just as some
acoustic differences are perceptually relevant and others are not, some perceptual cues
are phonologically relevant and others are not.

Is there a way to predict which perceptual cues are phonologically relevant and
which are not? To answer this question, we should perhaps consider factors that are
beyond the purely acoustic/auditory properties of cues. Releases and VC transitions are
different with respect to the edges of phonological categories they refer to or domains
they are part of. Releases are well ‘aligned’ with the right edge of a consonant (or pri-
mary gesture)12, and often with the right edge of a syllable, word, or utterance. VC tran-
sitions, and particularly palatalized transitions, are rather loosely ‘aligned’ with the left
dge of a consonant (or primary gesture), and are syllable-or word-internal. Given the
role of edges of phonological categories and domains in phonology [see Smith, 2004,
on edge effects in phonology], it is plausible that release cues are more cognitively
prominent and thus more relevant to higher–level phonological generalizations, com-
pared to VC transition cues. Further, stop releases are not limited to syllable-final posi-
tion, but are always present prevocally, prior to CV transitions. This may enhance
the cognitive prominence of releases, in the light of the universal preference for the CV
syllable type and the overall greater perceptual salience of CV transitions compared to
VC transitions [Ohala, 1990; Wright, 2001. (See Kochetov [2004] and Kavitskaya [in
press] on the role of CV and VC transitions in the perception of Russian palatalized
consonants.) In this respect, it is worth noting that Russian syllable-final stops are in
fact ‘used-to-be’ CV syllables [Kiparsky, 1979; Carlton, 1990].

Phonological relevance or irrelevance of acoustic/perceptual cues is not readily
accounted for in the hypothesis of licensing by cue, which assumes a phonetics/phonology
model where all phonetic knowledge of relative perceptibility is equally accessible to
phonological constraints (‘phonetics-in-phonology’ [Steriade, 1997, pp. 3, 4]). Perhaps
a more plausible interpretation of the ‘selective phonological attention’ to phonetic
cues can be provided if we assume a view of the grammar as multiple levels of pho-
netic/phonological generalizations that differ in their abstractness [Pierrehumbert et al.,
Phonetica 2006 – 145]
11 Conclusion

This study tested the hypothesis of licensing by cue [Steriade, 1997] by applying it to the distribution of the Russian plain/palatalized contrast in coronal stops in a number of syllable-final contexts: word finally and word internally before consonants (_#, _#k, _#n, _#s) and after vowels (a_, u_, i_). The hypothesis held that maintenance of the contrast /t/ versus /t̚/ in some of these contexts (_# and _#k, _#n and _#s) and its neutralization in other contexts (_#n and _#s) can be explained by the relative amount of information available about the contrast in each context. Acoustic and perceptual properties of the contrast were examined in a series of experiments involving word-internal and phonetically similar word boundary contexts. The results supported the hypothesis showing that relative salience of stop releases correlates with the word-internal patterns of neutralization of the contrast. However, the results also showed that robust acoustic and perceptual differences in VC transitions across vowel contexts had little to do with the attested patterns of neutralization. The findings are interpreted as evidence for a less direct, selective interaction between phonetic and phonological knowledge in the grammar. Finally, despite some apparent limitations, the hypothesis of licensing by cue [Steriade, 1997] provides an account of the facts of Russian neutralization that is explanatory in nature and empirically testable, and thus by far superior to many traditional phonological accounts.

Appendix

Table 11. Mean values of acoustic parameters for the minimal pair k[ak]/v/inverted/k[ak]/v/inverted produced by speaker 1 and used for experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Speaker 1</th>
<th>Speaker 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2 mid, Hz</td>
<td>1,289</td>
<td>1,379</td>
</tr>
<tr>
<td>F2 off, Hz</td>
<td>1,407</td>
<td>1,698</td>
</tr>
<tr>
<td>Closure dur., ms</td>
<td>67</td>
<td>50</td>
</tr>
<tr>
<td>Release dur., ms</td>
<td>21</td>
<td>60</td>
</tr>
<tr>
<td>Amp. diff., dB</td>
<td>28</td>
<td>9</td>
</tr>
<tr>
<td>Peak (part 1), Hz</td>
<td>3,639</td>
<td>5,785</td>
</tr>
<tr>
<td>Peak (part 2), Hz</td>
<td>5,350</td>
<td></td>
</tr>
<tr>
<td>C2 closure., ms</td>
<td>65</td>
<td>63</td>
</tr>
<tr>
<td>F2 on, Hz</td>
<td>1,651</td>
<td>1,869</td>
</tr>
</tbody>
</table>
Table 12.

<table>
<thead>
<tr>
<th></th>
<th><em>#</em></th>
<th><em>#</em></th>
<th><em>#</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>F₂ mid, Hz</td>
<td>1,143</td>
<td>1,427</td>
<td>1,233</td>
</tr>
</tbody>
</table>

Table 13.

<table>
<thead>
<tr>
<th></th>
<th><em>#</em></th>
<th><em>#</em></th>
<th><em>#</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>V duration, ms</td>
<td>132</td>
<td>131</td>
<td>118</td>
</tr>
</tbody>
</table>

Acknowledgments

References