Voice onset time across the generations

A cross-linguistic study of contact-induced change*

Naomi Nagy and Alexei Kochetov
University of Toronto

We investigate Voice Onset Time (VOT) of voiceless stops in conversational speech in a transitional bilingual context. We examine the speech of three generations of bilinguals whose Heritage Language (HL) is one of three European languages (Italian, Russian, or Ukrainian) and who also speak English. The data are extracted from recordings of sociolinguistic interviews conducted in Toronto that are contained in the Heritage Language Documentation Corpus (Nagy 2009). We examine word-initial /p, t, k/ in stressed syllables before /a/ and /o/ (~150 tokens per speaker), produced by 18 individuals representing three to five generations of speakers in each language. Unlike in English, voiceless stops in Italian, Russian, and Ukrainian are realized with a short lag VOT, defined as < 30 ms. Comparison of the HL patterns to previously published results on the VOT of monolingual speakers of these languages and monolingual speakers of English illustrates contact-induced influence: across the generations, the VOT of these speakers drifts away from the monolingual short lag toward the long lag of English for Russian and Ukrainian. Puzzlingly, the cross-generational change is (slightly) in the opposite direction for Italian. We discuss possible reasons for these different outcomes as well as contrasting them with the lack of cross-generational change found in analyses of pro-drop in the same corpus.

Keywords: language contact, Voice Onset Time (VOT), heritage language, Russian, Ukrainian, Italian, sociolinguistic variation, cross-generational change, Toronto

* This research is supported by SSHRC Standard Research Grant # 410-2009-2330. The research would not be possible without the generosity of our participants and the hard work of our research assistants, who are listed at http://projects.chass.utoronto.ca/ngn/HLVC. We would especially like to thank Vanessa Bertone, Natalia Harhaj, Melania Hrycyna, and Natalia Lapinskaya for their careful work measuring VOTs for this project.

1. Introduction

The study presented here forms part of the Heritage Language Variation and Change in Toronto Project (HLVC). This project compares the types of inter-speaker and cross-generational variation that occurs in different parts of the grammar of a range of heritage languages (HL) spoken in Toronto, a city where 44% (2.4 million people) count some language other than English as their mother tongue (Statistics Canada 2007a). The project compares large languages like Italian and Cantonese (each claiming >3% of the population as mother tongue speakers) to smaller languages like Ukrainian and Faetar. The long-range goals of the project are to better understand what cross-linguistic generalizations are possible about the types of features (or structures or rules or constraints) that are borrowed earlier and more often, and to understand the roles of social factors, on the individual and community levels, in these contact-induced changes.

Voice Onset Time (VOT) is defined as the duration (in seconds) of the interval between the release of a stop and the onset of vocal fold vibration. It is one of the cues that distinguishes voiced and voiceless obstruents in many languages (or lenis and fortis in some languages; Lisker & Abramson 1964). Voiceless stops in Italian, Russian, and Ukrainian are realized with a short lag VOT, defined as < .03 sec. English, in contrast, has a long lag VOT, defined as > .03 sec. Long VOT is often referred to as 'aspiration'. This difference makes VOT an excellent domain in which to explore sociolinguistic variation induced by language contact and better understand the path of linguistic drift in a transitional bilingual context. Our results reveal variation both within and across the heritage languages and correlations with indices of language contact. Because we are using conversational data, we expect (and find) bigger effect sizes than would be found in reading task data, given the greater effect sizes seen in less monitored speech styles (Labov 1972, among others).

Previous research shows that bilinguals (speakers who are fluent in two languages) can produce voiceless stop categories differently in each language. For example, simultaneous Canadian English-French bilinguals tend to realize English /p, t, k/ with a long lag VOT and the corresponding French stops with a short lag, as expected for both languages (Sundara et al. 2006, Fowler et al. 2008). Similar results were shown for native speakers of English who were advanced learners of French residing in France and native speakers of French who were advanced learners of English residing in the United States. Yet, the L1 and L2 categories appear to be cognitively linked and continuously influence one another. As a result, the bilingual production of stops in both L1 and L2 is different from that of monolinguals: the same English-French bilinguals were found to produce English stops with a VOT shorter than that of English monolinguals, and French stops with a

longer lag than that of French monolinguals (Fowler et al. 2008). We therefore expect to see that first generation speakers, whom we define as those that reached adulthood in the home country prior to immigrating to Toronto and have since spent at least 20 years in Toronto, will exhibit VOT patterns more similar to those of monolingual speakers of their L1, while second generation speakers (born in Toronto, or arriving before the age of six, with at least one first generation parent) and, to a greater extend, third generation speakers (children of second generation speakers) will have patterns more like monolingual English speakers.

Our research questions are, therefore:

- Do consistent patterns of change in VOT exist across and/or within languages?
- Are these patterns related to length of time that the family, or the community, has been in Toronto?
- Are these patterns related to (any aspects of) ethnic orientation?

2. Methods

In order to accurately analyze linguistic variation in order to detect change, it is important to compare apples to apples – in the HLVC project the same methods are used to select participants, collect speech samples and analyze data across the languages and across linguistic variables. We describe the languages and communities, the speakers and speech samples, and the methods of analysis.

2.1 Languages examined

The languages that form the HLVC project are summarized in Table 1. Of these, Italian, Russian, and Ukrainian are examined in this paper. Speakers of each of these languages have been present in Toronto for about a century, but they differ greatly in the ratio of mother tongue (MT) speakers to members of the ethnic population.[1] Almost half the Italian population reports Italian as the MT, while about one quarter of the Ukrainian population does so. The Russian figures are not directly comparable, as many people who are not ethnically Russian have Russian as their mother tongue (e.g., the large Russian-speaking Jewish community, many of whom have come to Canada via Israel and are listed separately from ethnic

1. "Mother tongue is defined by Statistics Canada as the first language learned at home in childhood and still understood" (Statistics Canada 2007b). By this definition, all speakers in our corpus are mother tongue speakers of their heritage language (a corpus selection criterion).

Russians, and Russian speakers from other countries that formed part of the ex-USSR). Necessarily, the group of people with Russian as MT, as a whole, will have a weaker link to Russian ethnic orientation in Toronto than languages that are more tightly connected to a particular ethnic background.

Table 1. Sketch of the languages included in the HLVC project with demographic statistics for the Greater Toronto Area

Language	# MT speakers	Ethnic population*	Date est.**	Place of origin of HLVC participants
Italian	186,000	< 466,000	1908	Calabria
Russian	65,000	≈ 59,000	1916	St. Petersburg & Moscow
Ukrainian	27,000	<< 122,000	1913	Lviv
Cantonese	170,000	< 537,000	1972	Hong Kong
Korean	49,000	= 55,000	1967	Seoul
Faetar	~100	= ~150	1950	Faeto & Celle St. Vito, Italy

* Population numbers from Statistics Canada (2007a, 2009).

** We have determined the date of establishment of the community to be the date of establishment of the first church in Toronto operating in the relevant language. As this does not apply for Faetar, which lacks institutional support and is not mentioned in census data, we use the earliest reported date of immigration among HLVC participants. Further details are available in Nagy (2011).

2.2 Speakers

We examine the speech of bilinguals whose Heritage Language (HL) is a European language (Italian, Russian, or Ukrainian) and who also speak English. Speakers defined as first generation were born and grew up in the homelands listed in Table 1 and have subsequently lived for at least twenty years in Toronto. Second generation speakers were born in Toronto (or arrived before the age of six) and have at least one parent who was born in the homeland. Third generation speakers are born in Toronto and have at least one parent qualifying as second generation. Thus, all speakers have been exposed to the geographic variety represented in the listed homelands.

Because these speakers live in a transitional bilingual context, we focus on inter-speaker variation. The data are drawn from the Heritage Language Documentation Corpus (Nagy 2009, 2011), consisting of sociolinguistic interviews conducted in Toronto with speakers in several generations of six heritage languages, stratified by age and sex. The data were collected 2009–2011. The interviews were conducted in the HL and produced about an hour of conversational speech from each participant, covering topics ranging from speaker's

upbringing and interests to their attitudes toward ethnic communities in Toronto. This approach allows us to describe naturalistic speech and requires us to carefully examine contextual effects. While our principal interest is examining variation in the HLs, a sister project is currently in progress to collect and analyze English speech from the same communities and using the same methods (Hoffman & Walker 2010). In anticipation of their results, and in light of the resource costs of recruiting speakers, interviewing, transcribing, and analyzing conversational speech (see next paragraph), we do not duplicate their efforts.

Speakers are recruited from the personal networks (where possible, and otherwise from extensions of these networks) of research team members who are themselves heritage language speakers who have grown up in Toronto. Participants are selected to participate if they self-define as "fluent enough to participate in an hour-long conversation in the heritage language." When complete, the corpus will contain samples of 40 such speakers from heritage languages, distributed across three generations, balanced for age and sex, and exhibiting a range of histories of language exposure and linguistic and cultural orientations. The 34 speakers analyzed here are listed in Table 2, represented by speaker codes which indicate their language, generation, sex, and age. The final letter in the codes below disambiguates speakers who would otherwise be coded identically and indicates the order of recording. It was not possible to avoid interactions between age and generation, due to limitations of available data, therefore age is not examined in this study.

Table 2. Speaker sample

Generation	Russian (N = 11)	Ukrainian (N = 12)	Italian (N = 11)
First	R1F55B	U1F85A	I1F71A
	R1M47A	U1M46A	I1F73A
	R1M56A	U1M85A	I1M62A
	R1M62D		I1M75A
Second	R2F17A	U2F54A	I2F44A
	R2F20A	U2F60A	I2F53A
	R2F50A	U2M56A	I2F57A
	R2M56B	U2M57A	I2M53A
Third	R3F25A	U3F26A	I3F21A
	R3F37A	U3F65A	I3F23A
	R3M56A	U3M24A	I3M22B
		U3M47A	
Fifth	–	U5F163A	–

2.3 Speech samples analyzed

In contrast to most previous studies of bilingual VOT based on experimental elicitations (read words or sentences) and examining stable sociolinguistic contexts, we investigate VOT in conversational speech.

To provide the data, speakers participate in three tasks beginning with a sociolinguistic interview (methodology defined in Labov 1984), a relaxed conversation conducted in the heritage language and digitally recorded. The goal of this task is to collect naturalistic in-group conversational speech. All interviews are conducted by heritage language speakers whose background is from the same region of origin as the participant (e.g., Calabria, for Italian participants, see Table 1). The interview data is transcribed orthographically in ELAN (Wittenburg et al. 2006) by native speakers of each HL. This method creates time-stamped transcriptions which are linked to the .wav file of the interview recording, making it possible to conduct visual examination of the audio file in Praat (Boersma & Weenink 2011) from within ELAN.

After this interview, an Ethnic Orientation Questionnaire (EOQ) is orally administered in the HL. Our questionnaire, adapted from Keefe & Padilla's (1987) study of Chicano ethnicity in the US, can be found online at http://projects.chass.utoronto.ca/ngn/pdf/HLVC/short_questionnaire_English.pdf. Participants are asked a range of questions about language use (their own and that of their family and friends), language attitude, and cultural orientation. These responses are recorded and coded. For each question, 0–2 points are assigned. A score of 0 indicates that the speaker (or her friends/family) uses or prefers English or orients toward Canada (depending on the question). A score of 2 indicates that the speaker uses or prefers the HL or orients toward the homeland. Mixed or ambiguous responses are scored 1 point. Scores of all answered questions are averaged to compute a numerical EOQ score for each participant. We also compute scores based on topic-based subsets of the questions (marked in the questionnaire at the above URL). A picture description task is also conducted, but that is not relevant to this study.

2.4 Acoustic analysis of VOT

In the ELAN transcription file, the first 25 instances of each segment of interest are marked, beginning at a time point 15 minutes after the beginning of the conversations, to decrease the possibility of examining unnatural speech as the participant gets used to the presence of the recorder and the interview situation. Marking within the ELAN transcription file permits us to recover any necessary context during all stages of analysis. Using the transcription, plus audio and visual cues from Praat, the tokens are segmented to mark the beginning and end of the

preceding segment (S), the closure (C), and the release (R) of the voiceless stop, and the following vowel (V). VOT was defined as the duration from the onset of the stop burst to the first zero-crossing of the first periodic wave of the following vowel. The following vowel's duration was measured as a means of controlling for speech rate differences among speakers. Figure 1 illustrates the mark-up for a Ukrainian token of word-initial /p/, in the word *pan* 'gentleman.'

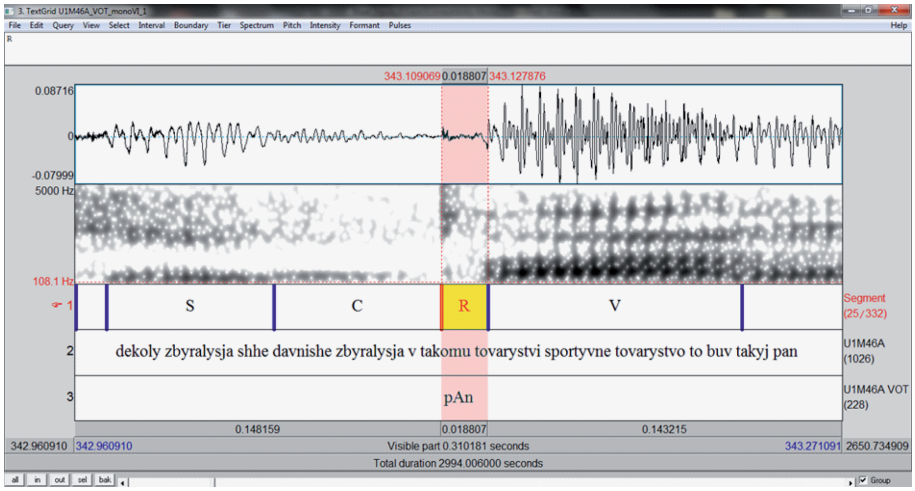

Figure 1. Intervals labeled in Praat for a Ukrainian token of word-initial /p/ by speaker U1M46A (.018 sec. release highlighted)

Only tokens that were clearly audible and free of speech errors and background noise were selected. A Praat script was run to extract the duration of each of these segments. We then conducted repeated measures ANOVAs to check for significant differences among consonants, between following vowels, and across generations, within each language. Correlations between mean VOT and EOQ scores among individuals were calculated.

Because of the possibility of great variation according to context, we restrict our examination to word-initial /p, t, k/ in stressed syllables before /a/ and /o/. Speech from 34 individuals representing three to five generations of speakers in each of the three HL languages is analyzed. While our original goal was to analyze 25 tokens for each consonant per speaker (resulting in 75 tokens per speaker), this was not always possible given differences in the duration of recording samples and inherent differences in the frequency of consonants in the three languages. For example, the Russian sample had considerably fewer tokens of /p/ than tokens of /t/ and /k/, while the Italian sample had fewer tokens of /t/ than /p/ and /k/. The vowel context was also unevenly distributed: for example, /pa/ was relatively rare

in Russian, while /to/ was rare in Italian. Actual numbers of analyzed tokens reflect this variation (as shown in Table 1 in the Appendix), giving on average 22 tokens of /p/, 27 of /t/, and 25 of /k/ per speaker and resulting in the total of 2,514 tokens (on average 74 tokens per speaker). On average, 50% of consonants were produced before /a/ and 50% before /o/.

In addition to the cross-generational comparisons, we compared the resulting VOT values to those previously reported for monolingual speakers of Russian (Ringen & Kulikov 2010, 2012), Calabrese Italian (Sorianello 1996), and Canadian English (Fowler et al. 2008). These comparisons should be taken with caution, however, as the VOT data in these studies come from read sentences. To the best of our knowledge, there are no comparable studies of VOT in conversational speech in the target languages; and we are not aware of *any* VOT studies of Ukrainian stops. We are, however, currently working with collaborators to produce comparable samples of Toronto English and Lviv Ukrainian, and to examine VOT in the Moscow spoken language portion of the Russian National Corpus.

3. Analysis

To ascertain the degree of contact-induced influence, we compare the HL patterns to those of monolingual speakers of each language. We report first on the patterns within each language, and then make cross-linguistic comparisons. We then consider the possibility of an effect of speech rate. Finally, we look at the connections between VOT and EOQ scores.

3.1 Cross-generational analysis of VOT, within each language

The Russian VOT data, summarized in Figure 2, generally illustrate the cross-generational trend that we expect: gradually lengthening VOT with each successive generation, suggesting drift toward English's longer lag VOT values. There is a significant difference between the mean measurements for the third generation vs. the second, although the difference between the first and second generations is not significant. The third generation's VOTs approach those reported for monolingual English speakers (in Montreal) by Fowler et al. (2008): /p/ 0.057 sec., /t/ 0.074 sec., /k/ 0.078 sec. The first and second generation VOTs, in contrast, are both within the range reported for Russian monolinguals in St. Petersburg by Ringen & Kulikov (2010, 2012): /p/ 0.018 sec., /t/ 0.020 sec., /k/ 0.038 sec. Recall that both the English and Russian monolingual data come from a reading task, which may show somewhat higher VOT values (cf. Lisker & Abramson 1964, Kessinger & Blumstein 1997).

The durations for /k/ are significantly longer than for /p/ and /t/, as expected for velar articulations. This same effect of place of articulation is observed in all the languages and samples reported here. Because this effect is well-established, we do not provide separate measurements for each place of articulation for the monolingual comparators. The curious reader will find these details in the cited studies.

By the third generation, another interesting trend emerges: the VOT duration for /t/ increases less than for /p/ and /k/. This suggests that (some) speakers may be developing two distinct categories: an alveolar English /t/ and a dental Russian /t/, resulting in the VOT of the second language exerting less influence on the first. In such a case, Flege's (1987) Equivalence Classification Principle would suggest that the Russian /t/ would be less influenced by English /t/ than would phonemes such as /p/ and /k/ which are more similarly articulated in the two languages (and stored as one category). This "dip" for /t/ is observable by contrasting the third generation columns of Figure 2 to those of the first and second generation.

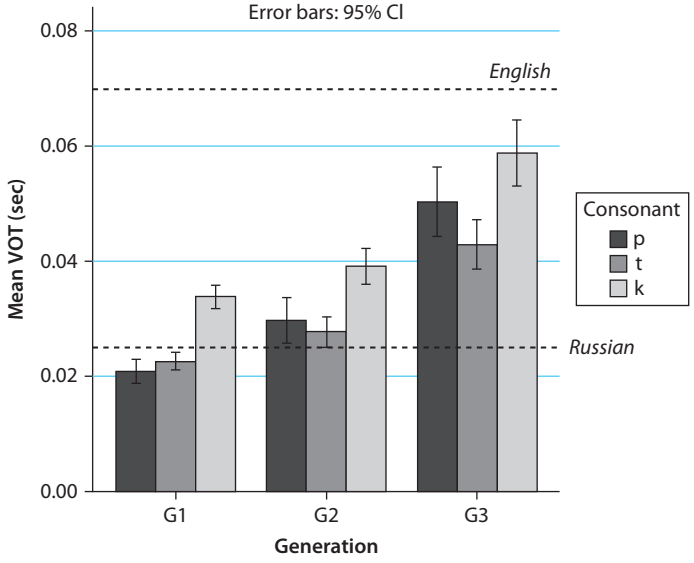

Figure 2. Russian VOT means for stops /p, t, k/ by generation (G1, G2, G3), compared to the Canadian English and homeland standards (based on the literature; see text for details)

The Ukrainian data (Figure 3) also shows the expected cross-generational changes, although in this case there is a significant increase from first to second generation, but not from second to third. We lack a homeland comparison data sample but believe that VOT in Ukrainian should be much like Russian. We also were

fortunate to record one fluent fifth generation Ukrainian speaker, who is included for comparison. She shows VOT measurements in the range expected for monolingual English speakers. The same "dip" for the /t/ emerges in the third generation Ukrainian data that was evident in the third generation Russian data.

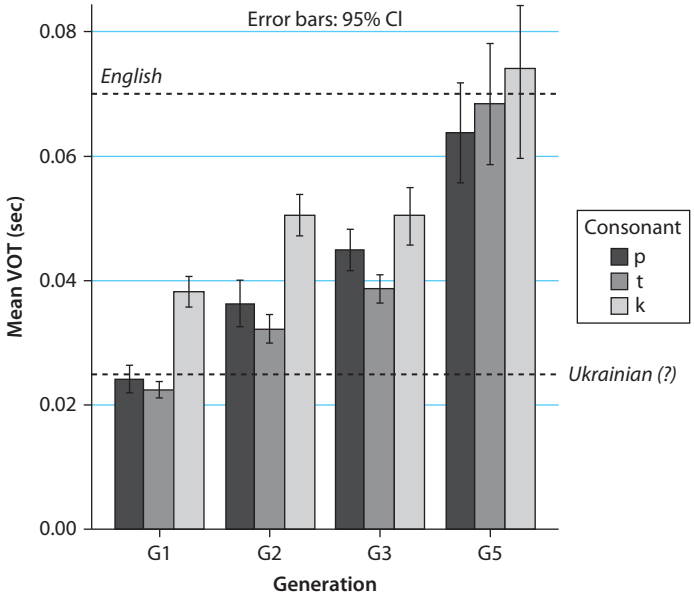

Figure 3. Ukrainian VOT means for stops /p, t, k/ by generation (G1, G2, G3, G5), compared to the Canadian English and presumed homeland standards

We turn next to the Italian speakers (Figure 4), who behave quite differently from the speakers of the two Slavic languages. For one thing, there is no cross-generational increase in VOT. In fact, there is a slight but significant decrease in VOT means from the first to the second generation. All generations show VOTs that are slightly longer than the means reported in Sorianello's (1996: 134) study of homeland Calabrese Italian speech, summarized in Table 3. Sorianello's means are for pre-tonic (not necessarily word-initial), intervocalic, non-phrase-final words produced by three speakers in a sentence reading task, and are the most similar comparison source available (being speakers from the homeland of the Toronto speakers and producing segments in similar linguistic contexts). Further, in Figure 4 we see no evidence of the "dip" for /t/ that the other two languages exhibit, though the similarity of values for /p/ and /t/ in Generation 3 (compared to Generations 1 and 2) suggests that it is almost established.

Table 3. Mean VOT values (in sec) in homeland Italian and Calabrese (adapted from Sorianello 1996: 134)

	Cosenza (Calabrese) dialect		Calabrese Regional Italian	
	mean	**standard deviation**	**mean**	**standard deviation**
/p/	.017	.03	.014	.07
/t/	.016	.04	.070	.04
/k/	.024	.06	.029	.06

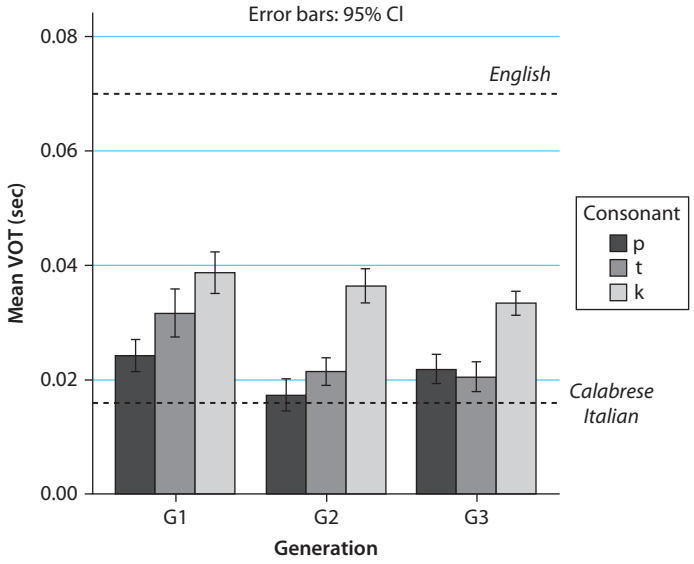

Figure 4. Italian VOT means for stops /p, t, k/ by generation (G1, G2, G3), compared to the Canadian English and homeland standards (based on the literature)

3.2 Comparison across languages

Figure 5 allows for comparison across the three languages, showing quite similar means for the first generation of all three languages, but different cross-generational trends of change in each: a significant change only between second and third generations in Russian but between first and second in Ukrainian, and a significant *decrease* from first to second generation for Italian. For comparison, we indicate English (from Fowler et al. 2008) and an averaged homeland value (from Ringen & Kulikov 2010, 2012 for Russian, Sorianello 1996 for Calabrese Italian). Note that these measurements come from sentence reading data, while the HL data is from conversational speech.

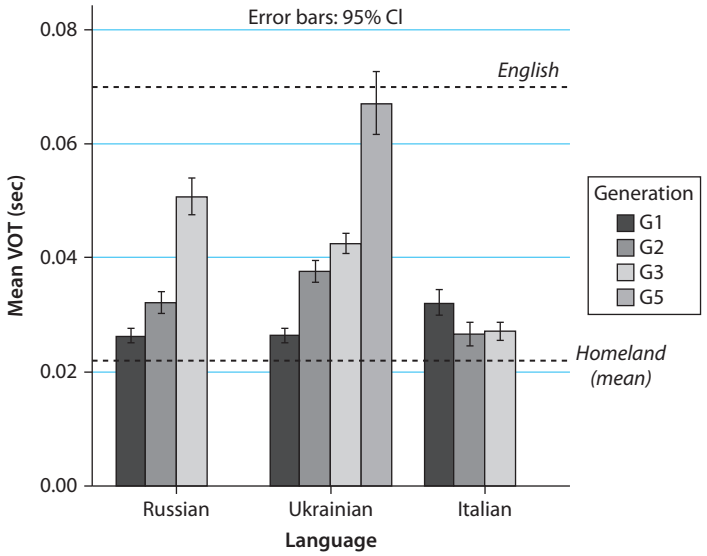

Figure 5. Cross-generation and cross-linguistic comparison of VOT means (from Hrycyna et al. 2011)

3.3 Potential effect of speech rate

We were concerned that differences in speech rate, likely slower speech in (some) third generation speakers than in (some) first generation speakers, might account for some apparent differences in VOT. That is, slower speech from the less fluent third generation speakers might result in longer VOT duration measurements for that group. In an attempt to control for that, we calculated the duration of the following vowel (which would also be longer in slower speech), in order to be able to examine the ratio of VOT to vowel length: a higher ratio would indicate a longer relative VOT, effectively normalizing for speech rate.

Figure 6 shows the vowel duration measurements for the three heritage languages. Note that for Italian speakers, the later generations have longer vowels, suggesting slower speech, as expected. Therefore, a cross-generational decrease in speech rate cannot account for the anomalous VOT results for Italian. If anything, the pattern illustrated in Figure 4 is an attenuation of the actual pattern: second and third generation VOTs would be even shorter if normalized according to the vowel durations which stand in for speech rate measures.

Russian patterns are similar to Italian, with respect to vowel duration: second and third generations have longer vowels than first generation, suggesting

slower speech. For Ukrainian, the fifth generation speaker has longer mean vowel duration than the other speakers, suggesting slower speech and a partial account for her much longer VOT measurements. However, the third generation shows a (non-significantly) lower mean vowel duration than the first and second generations, so there is no reason to question the VOT trend shown in Figure 3.

It is possible, however, that vowel duration *per se* does not correlate strongly with speech rate. Future analyses will include more detailed analysis of rate differences in our data, looking specifically at additional predictors such as numbers of syllables in the word or phonemes per time unit (cf. Cucchiarini et al. 2002, Kormos & Dénes 2004).

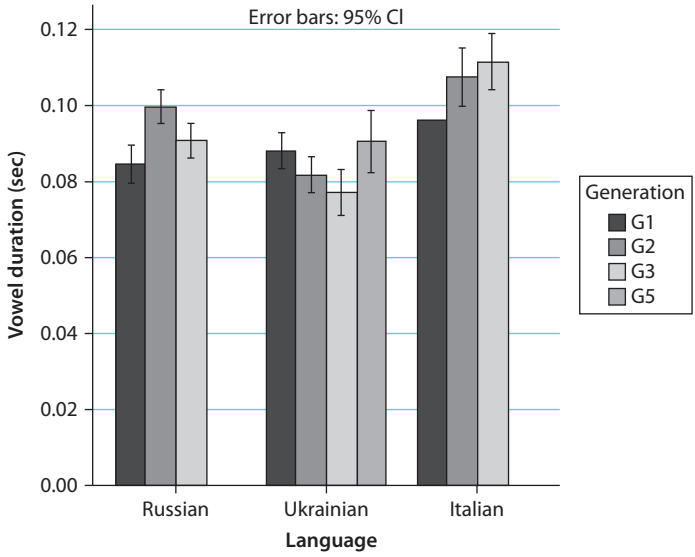

Figure 6. Mean duration of /a/ and /o/ vowels following the consonants examined for VOT, by generation

The reported VOT differences are confirmed statistically by repeated measures ANOVAs conducted within each language. We find a main effect for generation in Russian ($F(2,7) = 6.10$, $p < .05$) and Ukrainian ($F(2,7) = 12.01$, $p < .01$), but not Italian ($F(2,7) = 1.299$, $p = .33$). Bonferroni post-hoc tests show VOT to be significantly higher for the third generation compared to the first generation in Russian ($p < .05$; adjusted for multiple comparisons), higher for the fifth generation compared to the other generations in Ukrainian ($p < .01$-$.05$), and no generational effects for Italian. All languages have a significant effect for consonant, with /k/ significantly longer than other consonants, with p values ranging from 0.001 to 0.05.

3.4 Ethnic orientation effects

We turn next to the effect of Ethnic Orientation on VOT. It is important to note first that EOQ scores are highly related to generation in this sample: first generation speakers in all languages have higher overall EOQ scores than second generation, which, in turn are higher than third generation. (Reminder: Higher score means stronger orientation to the HL and/or homeland.) This trend is evident in the vertical dimension of Figure 7, where first generation speaker averages are higher in the graph, and later generations are lower. (There is one exception to the trend: Italian Generation 2 and 3 mean EOQ values are similar, but their ranks are switched.) The overall scores are necessarily related to generation, to a certain extent. For example, there is a question about where the participant and her parents were born, and its answer is fully dictated by our definition of generation. Therefore, we turn to a more fine-grained analysis of subparts of the EO questionnaire.

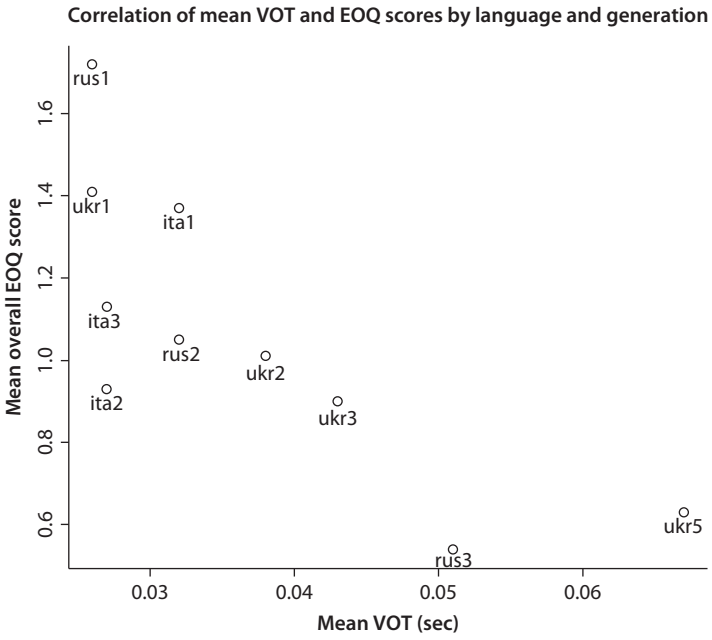

Figure 7. Average VOT on x-axis (higher values are more English-like) and average EOQ score on y-axis (lower values are more English-oriented) for each generation, each language

We look first at relationships within the EOQ subpart scores for the corpus as a whole. EOQ scores from 37 questions were divided into six indices of different aspects of the participants' linguistic and cultural behavior and attitudes. Table 4 illustrates the lack of correlation among these different subsets of EOQ responses, for 114 Italian, Russian, and Ukrainian speakers. Of all the pairs, only one is highly correlated ($|r| > .5$, *cf.* Cohen 1988: 83) and significant ($p < 0.05$): the correlation between cultural environment and language choice. This correlation holds across the languages as well as within each language. Given that these subsets are, therefore, with the one noted exception, independent measures, we might expect at least some of them to correlate to a linguistic pattern. That is, we might expect some correlation between how a participant speaks the HL and the speaker's attitude toward the HL and its speakers.

Table 4. Correlation across EOQ indices: Pearson's product-moment correlations (r)

	Language choices	Cultural environment	Language use	Cultural choices	Perceived discrimination
Ethnic ID	0.23	0.10	0.30	0.29	0.01
Language choices		0.81	0.21	0.20	−0.11
Cultural environment			0.25	0.12	−0.17
Language use				0.10	−0.02
Cultural choices					0.09

We turn now to look at correlations of these different subset measures of Ethnic Orientation and VOT scores, using data from the 22 speakers for whom we have both VOT data and responses to the EOQ. Table 5 provides Pearson's r values for each pairwise correlation. (Questions belonging to each subset are indicated in the online questionnaire cited above.) There are two high correlations in the data for all three language samples combined. The less a speaker uses their HL, the more their VOTs resemble English VOT values. And the lower the overall EOQ, the more English-like the VOT. These two correlations are also found when correlations are calculated just within the Russian and Ukrainian data sets. No high correlation between VOT and any EOQ score exists for Italian. We must be careful in considering the many high correlations in the Russian column, given that they are based on responses from only four speakers, as other speakers declined to answer too many of the questions in the questionnaire. The major finding here is that, where there is socially-demarcated variation (by generation) in a language's VOT scores, we find

that these values correlate to EOQ, and where there is no such social-marking, we find no correlation to EOQ (note the extremely low r-values for Italian).

Table 5. Correlation between EOQ indices and VOT: Pearson's product-moment correlations (r): High correlations (|r| > .5) in bold

EOQ topic	Italian (n = 7)	Russian (n = 4)	Ukrainian (n = 10)	3 languages combined (n = 21)
All questions	0.01	**−0.97**	**−0.74**	**−0.66**
Ethnic orientation	0.08	**−0.71**	−0.37	−0.31
Language choices	0.32	**−0.96**	−0.24	−0.29
Cultural environment	0.01	**−0.89**	0.01	−0.16
Language use	0.14	**−0.71**	**−0.60**	**−0.57**
Cultural choices	0.39	−0.34	**−0.51**	−0.29
Perceived discrimination	−0.34	**−0.95**	−0.20	−0.43

4. Summary

The answer to our first research question, whether consistent patterns of change in VOT exist across and/or within languages, is negative. In this study, we see the expected pattern of drift toward English VOT across generations for two of the three languages, but not for Italian. We see evidence of the development of separate phonemic categories for English and the HL (for /t/) in the same two languages, but again not for Italian.

Our second question is whether these patterns relate to the length of time that the family, or the community, has been in Toronto. Again, results are mixed. For two of the three languages, we see an effect of the length of time that the family has been in Toronto (reflected by cross-generational differences). The three language communities examined here have all been present in Toronto for just over a century, so we cannot consider this factor at the community level to be important. Although the beginning of immigration from the three countries is at about the same time point, the Italians immigrated at a much higher rate during the earlier periods of settlement, while most Russians and Ukrainians came more recently, and in a series of discontiguous waves. The long-time presence of a large Italian-speaking community may play a role in inhibiting English influence, but this remains to be determined once other variables are examined.

The answer to our third question, whether these patterns relate to (any aspects of) ethnic orientation, is affirmative. We see correlations between VOT of individuals and their scores on the EOQ as a whole and in the subsection related to language use, again with the caveat that Italians are an exception. No other subsets of the questionnaire play a consistent role.

5. Putting this study in its broader context

While it is true that in no case are English patterns entirely adapted in the HL spoken by later generation speakers in Toronto, we see a range of linguistic behaviors across these three languages. At this point, we can only speculate about reasons for the different behaviors. As noted at the outset, the goal of the HLVC Project is to gather sufficient data, using controlled methodology, to move beyond such speculation. As we are not there yet, here are some speculations that will become testable hypotheses in later stages of the project, once more linguistic variables are analyzed.

Because there is such a large, long-time Italian community in Toronto (10% of the city is ethnically Italian, and prior to 1991, it was by far the biggest source of immigrants to Toronto, Statistics Canada 2009), there is likely a great range of linguistic abilities. It may be that only the most fluent speakers come forward to volunteer for research projects like this one, and so we only see the very "best" speakers – those who maintain Italian as a quite different system from English. There is also a lot of institutional support for Italian in Toronto, meaning that many third generation speakers may get a great deal of input from the classroom, not just their family. This could account for the maintenance of homeland-like standards (Vanessa Bertone, p.c.). Additionally, Italians report discrimination against their group early in their history of migration to Toronto. This may have induced pressure early on to assimilate to English norms, possibly reducing the number of earlier immigrants who passed along their HL to future generations. This would decrease the pool of "not so careful" speakers from which we draw our Italian sample.

Ukrainian was an oppressed language in its homeland, outlawed from time to time under Russian or Polish hegemony. One of the goals of immigrants to Canada from Ukraine was to find a place where their culture and language could be practiced. As a result, there is a strong network of institutional support for Ukrainian, resulting in a variety of contexts beyond the family in which Ukrainian is spoken in Toronto (Melania Hrycyna, p.c.). 20% of people whose mother tongue is Ukrainian report that it is the language that they use most often at work (Statistics Canada 2010). This means that people born in Toronto may speak Ukrainian with a wide range of people outside the home, possibly accounting for shared norms between second and third generation speakers.

In contrast, Russian was never an oppressed language in its homeland. Russian immigrants may therefore not feel as strong a cultural pull to maintain their HL (Natalia Lapinskaya, p.c.). However, 24% of Russian mother tongue speakers report that Russian is the language they used most often at work (Statistics Canada 2010). Given the similarity in rates of usage of the HL at work, it is not obvious why, for Russians, the shift toward a more English-like VOT comes between the second and third generations, while for Ukrainian, it is between the first and second.

To highlight the importance of testing such hypotheses with a range of linguistic variables before drawing conclusions, we conclude by contrasting the VOT findings reported here to the patterns established for a morphosyntactic variable, pro-drop, or the variable surface presence of subject pronouns. Also using conversational data from the HLVC corpus, in some cases from the same speakers, Nagy et al. (2011) reported no significant differences between generations in either the rate of pronoun use or the factors conditioning pronoun presence in three HLs: Russian, Cantonese, and Italian. Although the three languages have different rates of usage (and represent partial, radical, and canonical pro-drop languages, respectively), they do not exhibit cross-generational differences. Considering the divergent outcomes of these two studies, it is evident that statements regarding the types of contact-induced change must be specific regarding both linguistic and social factors.

Abbreviations

ANOVA	Analysis of variance
C	Closure
EOQ	Ethnic Orientation Questionnaire
HL	Heritage language
HLVC	Heritage Language Variation and Change in Toronto Project
L1	First language
L2	Second language
MT	Mother tongue
VOT	Voice onset time
R	Release
S	Segment
SSHRC	Social Sciences and Humanities Research Council of Canada
V	Vowel

References

Boersma, P. & Weenink, D. 2011. Praat: Doing Phonetics by Computer. <www.praat.org> (4 October 2010).

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences,* 2nd edn. Hillsdale NJ: Lawrence Erlbaum Associates.

Cucchiarini, C., Strik, H. & Boves, L. 2002. Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America* 111: 2862–2873.

Flege, J. E. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47–65.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A. & Hallé, P. 2008. Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics* 36: 649–663.

Hoffman, M. & Walker, J. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22: 37–67.

Hrycyna, M., Lapinskaya, N., Kochetov, A. & Nagy, N. 2011. VOT drift in 3 generations of Heritage Language speakers in Toronto. *Canadian Acoustics* 39(3): 166–167.

Keefe, S. & Padilla, A. 1987. *Chicano Ethnicity*. Albuquerque NM: University of New Mexico Press.

Kessinger, R. H. & Blumstein, S. E. 1997. Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics* 25: 143–168.

Kormos, J. & Dénes, M. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32: 145–164.

Labov, W. 1972. *Sociolinguistic Patterns*. Philadelphia PA: University of Philadelphia Press.

Labov, W. 1984. Field methods of the project on linguistic change and variation. In *Language in Use*, J. Baugh & J. Sherzer (eds)., 28–53. Englewood Cliffs NJ: Prentice Hall.

Lisker, L. & Abramson, A. 1964. Cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20: 384–422.

Nagy, N. 2009. Heritage Language Variation and Change. <http://projects.chass.utoronto.ca/ngn/HLVC> (7 June 2012).

Nagy, N. 2011. A multilingual corpus to explore geographic variation. *Rassegna Italiana di Linguistica Applicata* 43(1–2): 65–84.

Nagy, N., Aghdasi, N., Denis, D. & Motut, A. 2011. Pro-drop in heritage languages: A cross-linguistic study of contact-induced change. *Penn Working Papers in Linguistics* 17(2): Article 16.

Ringen, C. & Kulikov, V. 2010. Voice Onset in Russian: 16th Annual Mid-continental Workshop on Phonology (MCWOP). Chicago IL, 29–31 October 2010. <myweb.uiowa.edu/vkulikov/MCWOP16_handout.pdf> (22 November 2011).

Ringen, C. & Kulikov, V. 2012. Voicing in Russian stops: Cross-linguistic implications. *Journal of Slavic Linguistics.*

Sorianello, P. 1996. Indici fonetici delle occlusive sorde nel cosentino. *Rivista Italiana di Dialettologia*: 123–159.

Statistics Canada. 2007a. 2006 Census of Canada: Population by Mother Tongue. <http://www40.statcan.ca/l01/cst01/demo12c-eng.htm> (5 December 2011).

Statistics Canada. 2007b. 2006 Census of Canada: Immigration, Citizenship, Language, Mobility and Migration. <http://www.statcan.gc.ca/daily-quotidien/071204/dq071204a-eng.htm> (22 November 2011).

Statistics Canada. 2009. 2006 Census of Canada: Immigrant Population by Place of Birth and Period of Immigration. <http://www40.statcan.ca/l01/cst01/demo24a-eng.htm> (22 November 2011).

Statistics Canada. 2010. 2006 Census of Canada: Language Used at Work. <http://www12.statcan.ca/census-recensement/2006/dp-pd/hlt/97-555a/pages/page.cfm?Lang=E&Geo=CMA&Code=535&Table=1&Data=CountO&StartRec=1&Sort=2&Display=All> (22 November 2011).

Sundara, M. L., Polka, L. & Baum, S. 2006. Production of coronal stops by simultaneously bilingual adults. *Bilingualism: Language and Cognition* 9: 97–114.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & Sloetjes, H. 2006. ELAN: A professional framework for multimodality research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.

Appendix

Table 1. Counts of tokens used in the study

Speaker/ Group	p	t	k	Total
R1F55B	3	25	27	55
R1M47A	24	25	24	73
R1M56A	18	25	25	68
R1M62D	23	25	24	72
R2F17A	25	25	25	75
R2F20A	11	25	18	54
R2F50A	25	25	25	75
R2M56B	19	24	24	67
R3F25A	22	25	21	68
R3F37A	8	23	25	56
R3M56A	20	22	24	66
Russian total	198	269	262	729
U1F85A	28	65	26	119
U1M46A	18	26	23	67
U1M85A	20	51	11	82
U2F54A	17	26	16	59
U2F60A	24	22	25	71
U2M56A	14	54	14	82
U2M57A	26	61	24	111
U3F26A	29	49	13	91
U3F65A	11	52	10	73
U3M24A	33	57	17	107
U3M47A	23	23	20	66
U5F16A	24	24	8	56
Ukrainian total	267	510	207	984
I1F71A	21	8	25	54
I1F73A	23	13	45	81
I1M62A	22	16	25	63
I1M75A	25	13	21	59
I2F44A	20	12	25	57
I2F53A	21	10	23	54
I2F57A	27	8	28	63
I2M53A	12	10	21	43
I3F21A	71	14	72	157
I3F23A	18	4	27	49
I3M22B	35	29	57	121
Italian total	295	137	369	801
All groups total	760	916	838	2514