CrossMark

Language Variation and Change, **26** (2014), 359–385.
© Cambridge University Press, 2014 0954-3945/14 $16.00
doi:10.1017/S0954394514000167

From interference to transfer in language contact: Variation in voice onset time

LUIZA NEWLIN-ŁUKOWICZ
New York University

ABSTRACT

This study examines cross-generational differences in the realization of an English phonological contrast by bilingual Polish Americans in New York City. I analyze the production of voice onset time (VOT) in underlying stops, as in *tin* and *den*, and stops derived from interdental fricatives, as in [t]in for *thin* and [d]en for *then*, in an English-only reading task. Generation one exhibits VOT "interference" for both stop types, with a bias toward interference for voiced stops. Generation two "transfers" Polish-like VOTs to derived stops. I argue that the cross-generational progression from the *global* effects of interference to the *focused* presence of transfer is filtered through L1 markedness and reflects speakers' growing sensitivity to L2 phonology and social considerations. The observed asymmetries in the distribution of interference/transfer are unaccountable by existing models of L2 acquisition and motivate a view of L1/L2 phonetic categories as governed by a variable grammar with access to phonological and social information.

Phonetic interference is a common characteristic of bilingual speech. Work in language contact and second language acquisition typically makes a distinction between the uncontrolled effects of L1 "interference" and the more controlled and often socially loaded occurrence of "transfer" of L1 features in L2 speech (Sankoff, 2002; Thomason, 2003). In the case of interference, influence from the L1 is seen to reflect the bilingual speaker's process of acquisition of the L2. In situations of language shift, L1 interference is expected to be lost in the speech of subsequent generations (Thomason & Kaufman, 1988; Van Coetsem, 1988). Transfer, or "reallocation" (Britain, 2002; Sharma & Sankaran, 2011), involves situations where generations of bilingual speakers retain some of that interference, instead of losing it. Thomason (2003:688) defined transfer as "innovations based on reinterpretation of source-language features by the speakers who implement the changes as well as the introduction of features

I am grateful to Lisa Davidson, Maria Gouskova, and Gregory Guy for valuable feedback on this work. I would also like to thank audiences at NWAV (New Ways of Analyzing Variation) 42 and the 88th Annual Meeting of the Linguistic Society of America for useful discussion of previous stages of this project. Lastly, the editors and anonymous reviewers have provided critical feedback that has led to many improvements to this paper. This research was supported by a National Science Foundation Doctoral Dissertation Research Improvement Grant (#BCS-1324916).

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

actually present in the source language." The changes that these L1-based features undergo may be structural or social (Britain, 2002), and in the latter case, typically involve the construction of an ethnic identity (e.g., Dubois & Horvath, 1998).

Despite the apparent prevalence of both interference and transfer, not much is known about which non-native features persist (or get transferred) across generations or why it happens. Thomason and Kaufman (1988) famously argued for the supremacy of social over linguistic factors in determining the outcomes of language contact. Yet variationist studies of second language acquisition have demonstrated that both linguistic and social factors determine the retention of interference (e.g., Adamson & Regan, 1991; Bayley & Preston, 1996; Preston, 1989; Sharma, 2005; Wolfram & Hatfield, 1984) and the acquisition of regional variants in the speech of ethnic groups (e.g., Drummond, 2012; Schleef, Meyerhoff, & Clark, 2011; Wolfram, Carter, & Moriello, 2004).

Studies of contact phonology in the speech of immigrant populations in the United States have repeatedly pointed to markedness as a linguistic predictor of which features are retained. Markedness is rather loosely defined and reflects both ease of learning and cross-linguistic frequency of occurrence (Thomason, 2010). Irrespective of their origin in L1 or L2, marked sounds are believed to be avoided in language contact situations. For example, interdental fricatives, universally difficult to acquire and typologically rare, are frequently stopped in the speech of bilinguals whose native phonologies lack these sounds, as in the case of the French-influenced Cajun English (Dubois & Horvath, 1998, 2000). However, this understanding of markedness provides only a partial explanation for the transfer effects observed. For instance, although it accounts for the presence of TH-stopping in the Cajun community, it does not explain the maintenance of other French-influenced variables, such as nasal realization of vowels, because oral vowels cannot be considered to be cross-linguistically marked.

This study examines cross-generational differences in the production of an English phonological contrast by bilingual Polish Americans living in New York City. At the segmental level, I analyze stop productions of interdental fricatives (TH-stopping), as in [t]in for *thin* and [d]en for *then.* TH-stopping is a regional feature of New York City English (NYCE), robustly present in the speech of Polish Americans (Newlin-Łukowicz, 2013). At the subsegmental level, I compare the realization of voice onset time (VOT) in stops derived from interdental fricatives and in underlying stops, as in *tin* and *den.* VOT, which represents the lag between stop release and the onset of voicing, has been identified as one of the main correlates of obstruent voicing in English and a common source of interference for bilingual speakers.

In an English-only reading task, the Polish-born first generation is found to exhibit a general pattern of VOT interference, with a bias toward interference with voiced stops. The New York City-born second generation transfers Polish-like VOTs to derived stops. I argue that the cross-generational progression from the *global* effects of interference to the *focused* presence of transfer is filtered through L1 voicing markedness and reflects speakers' growing sensitivity to L2

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

phonology and social differences between generations. The distribution of L1 interference and transfer emerges as variable and asymmetrical, and thus, cannot be accounted for by dominant models of L2 acquisition. The findings presented in this paper contribute to our understanding of bilingual phonology and motivate a view of L1/L2 phonetic categories as governed by a variable grammar.

VOT INTERFERENCE IN BILINGUAL SPEECH

Existing phonetic models of second language acquisition, such as the Perceptual Assimilation Model ([PAM] Best, 1995; Best & Tyler, 2007) or the Speech Learning Model ([SLM] Flege, 1995), regard the L1 as a type of "filter" that affects L2 acquisition. Both models posit that the development of an L2 phonetic category is contingent upon its perceived distance from existing L1 phones. The presence of interference then results from the initial misperception and the eventual misclassification of an L2 sound as belonging to an L1 category. Neither model provides an explicit metric of determining phonetic similarity or difference, but PAM posits that this process relies on the perception of articulatory features. For SLM, the successful acquisition of an L2 feature, with its concomitant acoustic or articulatory correlates, depends on perceptual (and phonotactic) factors that result from contrasting L1 and L2 feature specifications.

PAM and SLM are powerful models because they account for the most commonly observed outcomes of interference. For example, positing the presence of a single phonetic category for L1 and L2 sounds predicts "merged" VOTs (Chang, Yao, Haynes, & Rhodes, 2011), while assuming the formation of two separate sets of categories explains "compromise" and "polarized" VOTs (Flege & Eefting, 1987). The models' reliance on L1 and L2 phonological contrasts in phonetic category formation also explains divergent outcomes of interference for groups of bilinguals speaking different languages.

However, PAM and SLM do not allow for a variable presence of interference, whether at an individual or group level. Both models posit that, once established, L1 and L2 phonetic categories are used exclusively in their respective languages, unaffected by phonological or social factors. Sociolinguistic work, on the other hand, has uniformly pointed to the variable and systematic presence of interference in L2 speech (e.g., Adamson & Regan, 1991; Sharma, 2005), demonstrating that interference relies on multiple factors that can be modeled quantitatively. In addition, as models of phonetic interference, PAM and SLM do not account for cases of transfer. Yet, the maintenance of non-native features across generations can typically be traced back to interference in the speech of the first generation (e.g., Dubois & Horvath, 1998; Nagy & Kochetov, 2013; Sharma & Sankaran, 2011). Whereas sociolinguistic studies have commonly appealed to the social setting and cross-linguistic markedness as an explanation of transfer patterns, PAM and SLM predict that within-language markedness should matter as well.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

VOT: Polish vs. English

Polish-English bilinguals are susceptible to interference from VOT because Polish and English employ different VOT ranges. Like most other Slavic languages, Polish is a "voice" language in that it employs glottal pulsing to contrast prevoiced stops /b d g/ (negative VOT) with their voiceless counterparts /p t k/ (short-lag VOT). These two VOT categories are robust for Polish speakers in production (Keating, 1980) and perception (Keating, Mikoś, & Ganong, 1981) and show very little category overlap. Keating (1980) found Polish voiced stops to be consistently prevoiced, but the amount of prevoicing was larger in stop-initial isolated words (mean VOT for /d/ = –120 msec) than in stop-initial sentences (mean VOT for /d/ = –30 msec). By contrast, voiceless stops exhibited very little variation in VOT, and the mean VOT for /t/ was around 20 msec regardless of the task.

English, along with many other Germanic languages, can be classified as an "aspiration" language because it relies on the degree of openness of the glottis to distinguish unaspirated stops /b d g/ (short-lag VOT) from aspirated ones /p t k/ (long-lag VOT). English exhibits a very different VOT pattern from Polish, in that it contrasts two *positive* VOTs: short-lag VOT (small positive values, 0 to 10 msec) for unaspirated stops and long-lag VOT (large positive values, around 75 msec) for aspirated stops (Lisker & Abramson, 1964). Unlike in Polish, the two VOT categories show substantial overlap, so that there is no clear break between them. The degree of overlap is even more pronounced in the utterance-initial position, where English voiced stops lose their voicing (Flege & Brown, 1982). This situation again contrasts sharply with Polish, where prevoicing is most pronounced utterance-initially.

A number of studies have reported prevoicing for voiced stops in English, but variation across dialects and speakers is substantial. Jacewicz, Fox, and Lyle (2009) observed prevoicing in the utterance-medial position, when stops were flanked by (voiced) vowels, but only in North Carolina English, as opposed to Wisconsin English. Negative VOT has also been reported in post-pausal position, but here the occurrence of prevoicing differs across studies (e.g., Docherty, 1992; Flege & Brown, 1982; Lisker & Abramson, 1964; Simon, 2010).

VOICING: PHONOLOGICAL REPRESENTATION AND MARKEDNESS

These two contrastive voicing systems result in divergent within-language markedness relations. The short-lag VOT categories, "voiceless" in Polish and "unaspirated" in English, are believed to represent the unmarked members of the voicing pairs in the respective languages, because their production requires a *lack* of voicing or aspiration. By contrast, the marked members of the voicing pairs are the ones that require an extra articulation: prevoiced stops in Polish and aspirated stops in English.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

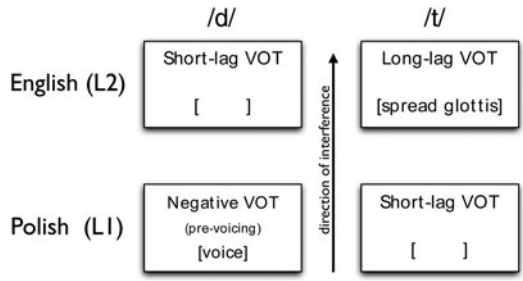


FIGURE 1. Contrasting Polish and English VOT systems.

To represent this distinction phonologically, I will follow the assumptions of "laryngeal realism," also known as "multiple feature theory" (Honeybone, 2005; Iverson & Salmons, 1995, 1999; Kager, van der Feest, Fikkert, Kerkhoff, & Zamuner, 2007). This model of the phonological representation of laryngeal contrasts is phonetics-based and draws on language-internal markedness distinctions. According to laryngeal realism, contrastive two-way systems, such as Polish and English, are distinguished by the presence of the privative feature [voice] and [spread glottis] on the marked member of each voicing pair. The unmarked members of each pair are left unspecified for either feature, resulting in the configuration in Figure 1.[1] Languages are commonly believed to manifest at most a three-way VOT contrast, in which case the feature [constricted glottis] is posited.

These conflicting featural specifications on the voicing pairs in Polish and English have predictions for the direction of interference. If L1 acts as a "filter" in L2 acquisition, and assuming that L1 and L2 categories coexist in shared phonological space (Flege, 1995), one could expect Polish /d/ to impose phonetic interference on English /d/ in the form of prevoicing because the former is specified for [voice] and the latter is not. Conversely, the Polish /t/ is not associated with any feature that could impose a specific realization on English /t/. Therefore, the realization of English /t/ in a monolingual-like manner (i.e., with long VOT) should depend on factors strictly related to the acquisition of a new L2 feature, [spread glottis].

Few studies have investigated the production of voiced stops in bilingual speech, as the focus has been placed on the voiceless series. However, voiced and voiceless stops typically behave differently in the speech of bilinguals whose L1 is a "voice" language and L2 an "aspiration" language. Specifically, L2 voiced stops always seem to be produced with elevated rates of prevoicing, whereas L2 voiceless stops may remain impervious to interference (and be produced with monolingual-like VOTs). For example, Punjabi-English bilinguals (Heselwood & McChrystal, 2000), French-English bilinguals (Sundara, 2005), and Italian-English bilinguals (MacKay, Flege, Piske, & Schirru, 2001) were all found to employ prevoicing more frequently than English monolinguals did.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

In what follows, I investigate whether within-language markedness distinctions can predict the direction of interference in language contact situations. PAM and SLM assume that L1 and L2 phonetic categories draw on the featural composition of phonological contrasts in their respective languages. Hence, I hypothesize that L2 phones will pose fewer acquisition problems (and trigger less interference) if they share the same phonological specification as their L1 counterparts than if their features mismatch. I explain the observed patterns of interference/transfer by assuming a variable grammar and by appealing to social factors beyond the scope of PAM or SLM.

TH-STOPPING

To test whether interference applies at uniform rates within the L2, I analyze VOTs in two contexts: underlying stops, as in *tin* and *den*, and stops produced through TH-stopping, as in [t]in for *thin* and [d]en for *then*. TH-stopping is typically defined as the "substitution" of stops [t d] for interdental fricatives [θ ð]. Yet, there are linguistic and social reasons to believe that speakers distinguish derived from underlying stops.

At the linguistic level, acoustic evidence suggests that speakers of American English differentiate underlying and derived voiced stops. Zhao (2010) found word-initial /d/ and stop-like realizations of /ð/ to diverge on several acoustic measures of the burst. The observed differences systematically confirmed that the place of articulation is alveolar for /d/ and dental for stop-like /ð/, corroborating the common impression that stopped /ð/ and /θ/ retain the place of articulation of the original fricatives (e.g., Labov 1966). Yet, Zhao's findings are not informative about the realization of /ð/ in dialects where TH-stopping is a regional feature, as speakers of these dialects, including New Yorkers, were purposely excluded, under the assumption that they would not differentiate the two stops. The stopped realizations in her study then reflect sporadic articulatory modifications, rather than systematic sociolinguistic variants.

Different places of articulation for derived and underlying stops should also be reflected in their VOTs, because VOT is expected to increase as the place of articulation becomes more posterior. For example, Lisker and Abramson (1964) report the shortest mean VOTs for bilabial stops (/b/: 1 msec, /p/: 58 msec), intermediate VOTs for alveolar stops (/d/: 5 msec, /t/: 70 msec), and the largest VOTs for velar stops (/g/: 21 msec, /k/: 80 msec). Similarly, Sundara (2005) compared the acoustics of coronal stops /t d/ in Canadian English and Canadian French, where these stops are described as alveolar and dental, respectively. She found that VOTs for coronal stops are longer in Canadian English (VOT for /t/ = 60 msec, /d/ = 16 msec) than in Canadian French (VOT for /t/ = 20 msec, /d/ = –82 msec), presumably reflecting differences in the place of articulation of these two stops.

At the social level, derived stops could be distinguished from their underlying counterparts because of the social meaning that TH-stopping may carry as a

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

common sociolinguistic variable (Wells, 1982). In NYCE, TH-stopping is a stigmatized dialect feature, associated with class-based, stylistic, and ethnic differentiation (Labov, 1966). Its origins have been anecdotally attributed to the non-native speech of the first immigrants to the area, such as the Irish, Italians, and Poles (Babbitt, 1896; Labov, 1966). This hypothesis is largely uncontroversial as substitutions of interdental fricatives are common for L2 learners whose native phonologies lack interdental fricatives (Lombardi, 2003). Labov (1966) estimated TH-stopping to be in stable variation, but it may be receding from contemporary NYCE, just like other stigmatized local features are (Becker, 2010). Yet, TH-stopping remains robust in the speech of Polish New Yorkers, patterning in a way that is consistent with L1 interference but also an ethnic marker (Newlin-Łukowicz, 2013).

THE STUDY

Participants

The present study was conducted as part of a larger research project on language variation among Polish New Yorkers. This larger project integrates variationist, quantitative methodologies with a qualitative approach that draws on ethnographic research with this group (as in neighborhood studies, e.g., Hall-Lew, 2009, and peer group studies, e.g., Mendoza-Denton, 2002). As the present study crosscuts a number of neighborhoods and communities of practice in New York City, the goal of the ethnography was to obtain relevant social information about Poles and Polish Americans residing in New York City. Ethnography involved participant observation of language use in everyday activities in Greenpoint, Brooklyn, and in cultural activities in greater New York City. Ethnographic observations were conducted over the span of three years (fall 2009 to fall 2012), culminating in data collection in the spring of 2013.

Thirty-five Polish Americans were sampled to participate in the present study. Participants were bilingual children of Polish parents who immigrated to New York City between 1970 and 1990. Participants classified as generation one ($n = 15$; 8 women, 7 men) immigrated to New York City and started learning English between the ages of 8 and 15.[2] Those classified as generation two ($n = 20$; 10 women, 10 men) were born in the city or immigrated and started learning English before age 5. All participants were close in age (19 to 35; one speaker was 46) and had similar social class backgrounds: they were raised in working-class homes and neighborhoods and held college degrees or were enrolled in college at the time of the study.

Participants' bilingual abilities were not formally tested. Instead, the author was able to verify participants' ability to speak Polish through ethnographic observation, as well as in additional tasks, such as sociolinguistic interviews and use-of-Polish surveys. Almost all second-generation speakers underwent some formal education in Polish, having attended various Polish supplementary schools in New York City for up to 13 years, and for an average of 5 years.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

All second-generation speakers reported having acquired the two languages sequentially, with Polish being spoken from birth and English from about age 5. They described themselves as English-dominant in sociolinguistic interviews, and many stressed having difficulties reading and writing in Polish. In the use-of-Polish survey, all second-generation speakers reported speaking only Polish to their extended families. Half of them also reported speaking exclusively Polish to their parents, whereas the other half spoke both Polish and English. All but one second-generation speaker reported using exclusively English or code-switching between Polish and English in interactions with their siblings and Polish American friends.

All first-generation speakers in this study underwent some formal schooling in Poland (from two to eight years), and some of them attended Polish supplementary school in New York City for another one to five years. These participants typically spoke little English before moving to New York City, and some still retained traces of a non-native accent. Many expressed difficulties communicating in Polish, much like the second-generation speakers did. Their language choices also mirrored those of the other group, with Polish being spoken at home and English or both languages elsewhere. In the home, first-generation speakers reported speaking Polish to parents as well as siblings.

In addition, 10 white non-Polish New Yorkers (7 women, 3 men) were recorded to obtain baseline data on TH-stopping and VOT for the New York City dialect. These speakers were all born in New York City, spoke English at home, and reported various European ancestries (i.e., Irish, Italian, German, Danish, and Eastern European). Because of impressionistic accounts that suggest that TH-stopping is diminishing in New York City, half of the speakers matched the sample of Polish Americans in age (20 to 41 years), and half of them were older (51 to 61 years). All speakers were monolingual in English, but had some exposure to other languages in schools (e.g., Spanish, French) or within the family (e.g., Hebrew).

Materials and procedure

Participants were asked to read three short stories loosely based on stories told by American stand-up comedians for Reader's Digest (see the Appendix). The short stories were heavily modified to accommodate a large number of tokens of word-initial underlying coronal stops /t d/ and interdental fricatives /θ ð/ and were designed with the intention of making participants produce TH-stopping. Newlin-Łukowicz (2013) found that, relative to the sociolinguistic interview, Polish Americans' stopping rates dropped drastically in a word list task, but remained high in a short story task. Therefore, to prevent participants from becoming overtly aware of the linguistic nature of the task, the short stories in this study were kept intentionally light-hearted. As such, they were designed to resemble spoken, rather than written, texts and included a number of incomplete sentences, exclamations, and expressions commonly encountered in spoken speech.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

Using short stories rather than spontaneous speech allowed for the control of a number of linguistic factors that are crucial in comparing VOT categories. The linguistic context is particularly essential in distinguishing prevoicing from a continuation of voicing from a preceding voiced segment. I thus controlled for the position of interdental fricatives and stops within the utterance (initial, medial, final), the segment that preceded them (pause, vowel, voiceless stop, voiceless fricative), and the words in which they appeared. No preceding voiced sounds were included, except for vowels, which are known to condition prevoicing for American English speakers (Jacewicz et al., 2009), and thus provide a point of comparison for stops produced after voiceless segments.

The distribution of word-initial interdental fricatives in the English lexicon is such that the voiced one is largely limited to function words, and the voiceless one to lexical items. Accordingly, this asymmetry is reflected in the short stories. Tokens of /ð/ are represented by the function words *that*, *this*, and *there*. These function words were chosen based on their overall high TH-stopping rates as compared to other function words (Newlin-Łukowicz, 2013), as well as their ability to appear in various positions within the utterance. For example, *this* can be utterance-initial ("*This* is when I knew she was a crazy person"), utterance-medial ("Let's kick *this* party off with a bang"), or utterance-final ("He looks deep into my eyes, like *this*"). Tokens in the utterance-medial and utterance-final positions were preceded by vowels, voiceless stops, or voiceless fricatives (three tokens of each per position). To boost the number of /ð/ tokens, each function word appeared three times in each utterance position.[3] I included a larger number of tokens of the voiced interdental fricative than of the other phonemes as TH-stopping was not expected to apply to all cases of /ð/. Tables 1 and 2 contrast the targeted and actual number of tokens. The actual number of tokens was always lower than the targeted number.[4] For derived stops, it was due to the variable nature of TH-stopping. For underlying stops, it had to do with the absence of a clear burst for some stop tokens, which excluded them from analysis.

The distribution of /θ/ in the short stories mirrored that of /t d/, due to similar lexical properties. For /θ/, as well as /t d/, six lexical items appeared in each utterance position, as there are fewer limitations on the distribution of lexical items within sentences than there are for function words. Therefore, in each utterance position, /θ/ was represented by one token of a monosyllabic or

TABLE 1. *Targeted number of tokens per stop type and utterance position, for the entire sample of Polish Americans and per speaker (in parentheses)*

Sound/Utterance Position	Initial	Medial	Final	Total
Underlying /ð/	315 (9)	315 (9)	315 (9)	945 (27)
Underlying /θ/	210 (6)	210 (6)	210 (6)	630 (18)
Underlying /d/	210 (6)	210 (6)	210 (6)	630 (18)
Underlying /t/	210 (6)	210 (6)	210 (6)	630 (18)
Total	945 (27)	945 (27)	945 (27)	2,835 (81)

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

TABLE 2. *Actual number of tokens produced by Polish Americans*

Sound/Utterance Position	Initial	Medial	Final	Total
Underlying /d/	210	173	180	563
Underlying /t/	208	165	188	561
Derived /d/	165	135	92	392
Derived /t/	81	103	59	243
Total	664	576	519	1,759

disyllabic word: *think*, *thank*, *thousand*, *thing*, *thrilled,* and *three*. Underlying coronal stops were, similarly, represented by tokens of monosyllabic and disyllabic words: *deep*, *Deb*, *David*, *days*, *doors*, and *dark* for /d/, and *Tina*, *tell*, *take*, *tables*, *talk*, and *Tom* for /t/. In the utterance-medial and utterance-final positions, one third of tokens of /θ t d/ were preceded by a vowel, a voiceless stop, and a voiceless fricative. For underlying stops, I additionally controlled for the following vowel and included tokens of high /i/, mid /ɛ, eɪ/, and low /ɑ, ɔ/.

Recordings were made in a sound attenuated room, using a Zoom H4N digital recorder, set at a sampling rate of 44.1 kHz. Each participant was recorded individually, after having completed a sociolinguistic interview and a use-of-Polish survey. Participants wore an Audio-Technica ATR3350 lavalier microphone and were instructed to read the short stories out loud and at their regular reading pace.

Acoustic analysis

All tokens of interdental fricatives were examined spectrographically in Praat (Boersma & Weenink, 2014). Underlying interdental fricatives were coded as fricatives, stops, or affricates, following the criteria laid out in Table 3. These coding criteria are fairly conservative, as I take the presence of a burst (in addition to the presence of closure) to be a requirement for the identification of stops. Stops identified this way were later subjected to the VOT analysis.

VOT was measured in word-initial underlying stops (/t/, e.g., *tin*, total $n = 724$; /d/, e.g., *den*, total $n = 733$) and stops derived from interdental fricatives ([t], e.g., [t]in for *thin,* total $n = 290$; [d], e.g., [d]en for *then,* total $n = 475$). See Table 5 for the breakdown of tokens per speaker group. Using TextGrids in Praat, I segmented VOT boundaries for each stop token, and their durations were

TABLE 3. *Acoustic criteria for coding*

Realization	Criteria
Fricative [θ ð]	Presence of frication
Stop [t d]	Presence of closure and burst, no frication
Affricate [tθ dð]	Presence of closure and burst, followed by frication

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

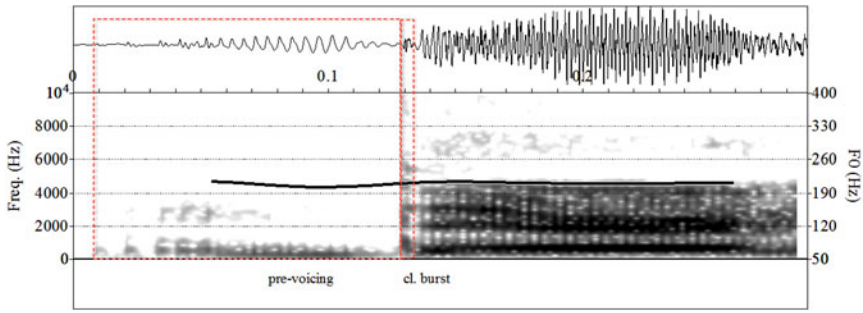


FIGURE 2. Prevoicing and closure burst in a stopped realization of *there* by a second-generation speaker.

extracted using a script. For prevoiced stops, VOT was measured from the onset of voicing (determined by the presence of the voice bar and periodic waveform preceding the burst) up until the burst (see Figure 2). For stops that did not show prevoicing, VOT was measured from the burst to the start of periodic waveform in the vowel. Measurements for prevoiced stops were recorded with negative values; the rest were positive.

Statistical analysis

VOTs were analyzed through mixed-effects regression modeling in R (R Development Core Team, 2011), using the *lmer()* function of the lme4 package (Bates & Maechler, 2009). Separate models were built for generation one, generation two, and non-Polish New Yorkers. The models tested included the following fixed effects: STOP TYPE ([t], [d], /t/, /d/), UTTERANCE POSITION (initial, medial, final), PRECEDING ENVIRONMENT (pause, vowel, voiceless fricative, voiceless stop), and GENDER (male, female; not tested for the non-Polish group). SPEAKER and TOKEN were included as random slopes to account for within-speaker and between-speaker variability in the VOTs produced for the four categories of STOP TYPE. The models reported in the following section constituted the best fit arrived at through a step-wise evaluation process whereby nonsignificant fixed effects were discarded (Baayen, 2008:182–185). The inclusion of each fixed effect was determined by an analysis of variance model comparison, based on likelihood ratio tests. Linear mixed-effects modeling was followed by post hoc Tukey tests (the *glht* function in the *multcomp* package) that determined which levels of a predictor were significantly different from each other. This method is conservative as it employs Bonferroni adjusted *p* values.

RESULTS

All speakers were found to produce TH-stopping while reading the short stories, but some did so more often than others. Non-Polish New Yorkers produced

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

TABLE 4. *TH-stopping rates for Polish Americans, divided by generation and gender*

Stop		Generation One			Generation Two		Total *n*
[d]: [d]en for *then*	Female	115/203 (56%)	197/376 (52%)	Female	109/281 (39%)	195/507 (38.5%)	392
	Male	82/173 (47%)		Male	86/226 (38%)		
[t]: [t]in for *thin*	Female	58/143 (40.5%)	109/260 (42%)	Female	70/197 (35.5%)	134/363 (37%)	243
	Male	51/117 (43.5%)		Male	64/166 (38.5%)		
Total *n*	306			329			635

stopping at an average rate of 29% and showed a slight preference for stopping of the voiced fricative (30%) relative to the voiceless one (26%). Older and younger non-Polish New Yorkers produced TH-stopping at the same average rate, which suggests that TH-stopping remains a stable feature of the dialect. Polish Americans exhibited rates of stopping that were considerably higher than these baseline values (see Table 4). Generation one showed the highest stopping rates: 52% for /ð/ and 42% for /θ/. Generation two produced stopping at a rate that was intermediate between that for generation one and non-Polish New Yorkers (i.e., around 38% for both /ð/ and /θ/).

All speakers produced a wide range of VOT values. Figure 3 presents density plots of VOTs per stop type for the two generations of Polish Americans and non-Polish New Yorkers. The two voiceless stops are realized the same by all groups: tokens of underlying /t/ are overwhelmingly long-lag, and tokens of derived [t] are mainly short-lag. Differences between speaker groups only emerge for the two voiced stops. Specifically, density plots for underlying /d/ and derived [d] are bimodal for Polish Americans, as all voiced stops are

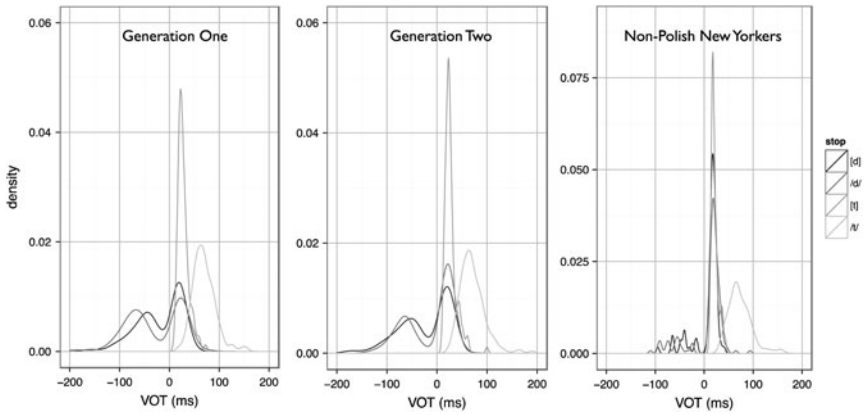


FIGURE 3. Density plots of VOTs produced by the three speaker groups.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

variably produced with negative and short-lag VOTs. Whereas generation one displays a comparable ratio of prevoiced to short-lag tokens for both stops (55% of prevoicing for /d/ and 50% for [d]), generation two trends toward English-like, short-lag VOTs for underlying /d/, where prevoicing is at 40%, compared to 49% for derived [d]. By contrast, non-Polish New Yorkers equally prevoice derived and underlying stops and do so at the much lower rate of 17%.

Linear mixed-effects modeling

Modeling revealed that VOT values are largely affected by the same factors for all three groups, with significant effects of STOP TYPE, PRECEDING ENVIRONMENT, UTTERANCE POSITION, and a significant interaction OF STOP TYPE and PRECEDING ENVIRONMENT. Second-generation New Yorkers additionally show an effect of GENDER. Table 5 breaks down the number of tokens speakers produced for different factor levels, and Table 6 summarizes regression results. In Table 6, the intercept represents the baseline to which the remaining factor levels are compared. In this case, it stands for the estimated VOT for underlying /d/ in utterance-final position and following a voiceless fricative. The estimates for the remaining factor levels indicate the amount by which the mean is predicted to increase (if the estimate is positive) or decrease (if the estimate is negative), relative to the estimate for the intercept. For ease of interpretation, factor levels are marked in **bold** if significant (*t* value exceeds –2 or 2).

Models for all three groups reveal that VOTs vary according to STOP TYPE. This result is expected, as English stops are known to fall into two VOT categories: short-lag and long-lag. Table 6 confirms that this distinction is present for all three groups: the mean VOTs for underlying /t/ are about 50 msec longer than the mean VOTs for underlying /d/ (the intercept) are, suggesting that the former stops type is aspirated. For non-Polish New Yorkers, all stops fall into this English-like distinction. These speakers contrast long-lag VOT for /t/, as in *tin*,

TABLE 5. *Factor levels used in the regression analyses, along with the number of tokens*

		Generation One	Generation Two	Non-Polish
STOP TYPE	Underlying /d/	238	325	170
	Underlying /t/	233	328	163
	Derived [d]	197	195	83
	Derived [t]	109	134	47
PRECEDING ENVIRONMENT	Pause	292	346	165
	Vowel	223	265	119
	Voiceless stop	143	194	98
	Voiceless fricative	119	177	81
UTTERANCE POSITION	Initial	300	364	171
	Medial	244	332	148
	Final	233	286	144
GENDER	Women	427	535	311
	Men	350	447	152
TOTAL		777	982	463

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

TABLE 6. *Fixed-effect coefficients in a mixed-effects model fitted to VOT values*

		Generation One		Generation Two		Non-Polish	
		Estimate	*t*	Estimate	*t*	Estimate	*t*
INTERCEPT	Underlying /d/, voiceless fricative, utterance-final	14	1.2	13	1.9	16	2.2
STOP TYPE:	Derived [d]	**−25**	**−2.1**	**−40**	**−3.9**	1.8	.18
PRECEDING	Derived [t]	10.1	.8	8	.9	4.7	.4
ENVIRONMENT	Underlying /t/	**54**	**4.25**	**44.7**	**4.5**	**50.5**	**5.5**
	Pause	**−71.3**	**−6.8**	**−48.2**	**−4.4**	**−33.6**	**−3.35**
	Vowel	**−55.4**	**−6.2**	**−67.2**	**−10.2**	**−25.4**	**−3.7**
	Voiceless stop	1.5	.15	−5	−.7	14.5	1.9
STOP TYPE:	[d]: Pause	**35.6**	**3.4**	**26**	**2.9**	−7	−.7
PRECEDING	[t]: Pause	**55.2**	**4.6**	**24**	**2.6**	8	.7
ENVIRONMENT	/t/: Pause	**50**	**4.5**	**33.3**	**4**	15.4	1.8
	[d]: Voiceless stop	11.3	.9	18.6	1.8	−14	−1.2
	[t]: Voiceless stop	6.4	.4	11.1	.9	−9.2	−.7
	/t/: Voiceless stop	.85	.06	16.7	1.6	−13.2	−1.2
	[d]: Vowel	**25.7**	**2.1**	**50**	**4.8**	−10.4	−.8
	[t]: Vowel	**62.4**	**4.9**	**66.8**	**6.6**	**28.2**	**2.1**
	/t/: Vowel	**54**	**4.6**	**72**	**7.7**	**27**	**2.75**
UTTERANCE POSITION	Initial	**21.7**	**3.2**	**28.5**	**3.15**	**31**	**3.9**
	Medial	−2.1	−.65	1	.4	−5.3	−1.8
GENDER	Women	N/A		**5.6**	**2.1**	N/A	

Note: **Bold** values are statistically significant.

with short-lag VOT for everything else ($p < .001$): the two voiced stops and derived [t]. This means that they do not distinguish [d]en for *then* from *den*, producing both with short-lag VOT, but they do distinguish [t]in for *thin* from *tin*, deaspirating the former. Figure 4 illustrates this two-way contrast.

The two generations of Polish Americans depart from this pattern. First, they differentiate derived [t] from the voiced stops and realize it with a VOT that is intermediate between that of underlying /t/ and /d/. Although the *t*-test comparison between derived [t] and underlying /d/ is not significant ($p = .9$), these two stop types are distinguished in their VOT distributions (see Figure 3):

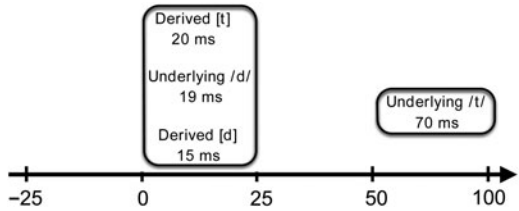


FIGURE 4. VOT means for stop categories produced by non-Polish white New Yorkers.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

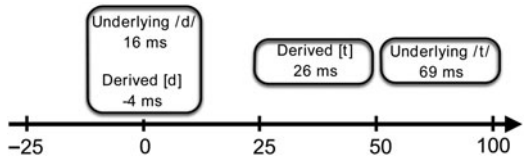


FIGURE 5. VOT means for stop categories produced by first-generation Polish Americans.

derived [t] is never realized with a negative VOT. Polish Americans thus produce an extra VOT contrast that is absent from the non-Polish group. Second, Polish Americans employ considerably more prevoicing than the non-Polish group does, which affects their realization of voiced stops. Generation one produces a comparable ratio of prevoiced and short-lag tokens for underlying /d/ and derived [d] ($p = .13$), but it shows a slight preference for short-lag tokens with underlying /d/, which is reflected in the positive mean for this stop. The first generation then produces a three-way VOT contrast that distinguishes prevoiced [d] and /d/, short-lag [t], and long-lag /t/, as illustrated in Figure 5. The second generation polarizes the contrast between the two voiced stops, restricting prevoicing to the derived context. As such, the second generation adds a fourth dimension to the VOT distinctions in English stops, as it produces negative VOT for [d], short-lag VOT for /d/, (longer) short-lag VOT for [t], and a long-lag VOT for /t/ (see Figure 6). The *t*-test comparisons of all stop types are highly significant and yield $p < .001$ (with the exception of the /d/-[t] comparison explained earlier). This four-way VOT distinction is unlike any other VOT contrast that has been reported, as (monolingual) language varieties are believed to distinguish no more than three VOT categories (e.g., Lisker & Abramson, 1964).

PRECEDING ENVIRONMENT also emerged as a predictor of VOT. Table 6 shows that underlying /d/ (the intercept) differentially favors short-lag and negative VOT. For all three groups of speakers, short-lag VOT is more likely to appear if underlying /d/ is preceded by a voiceless fricative (intercept; mean VOT between 13 and 16 msec) or a voiceless stop (mean VOT between 8 and 30 msec), whereas prevoicing is conditioned by a preceding vowel (mean VOT between –54 and –9 msec) or pause (mean VOT between –57 and –17 msec). The postvocalic position is known to produce prevoicing for American English (Jacewicz et al., 2009) and likely represents a continuation of voicing from a preceding voiced segment. The preference for negative VOT in the postpausal context is more unusual and possibly points to emphatic pronunciations. The postpausal context included

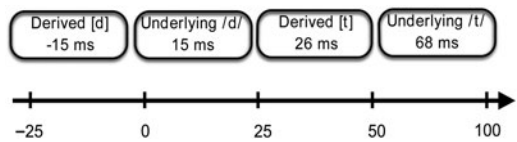


FIGURE 6. VOT means for stop categories produced by second-generation Polish Americans.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

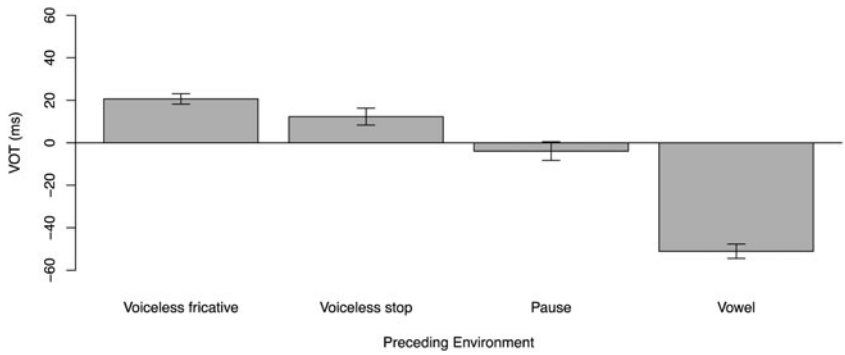


FIGURE 7. Mean VOT for underlying /d/ depending on preceding environment for second-generation Polish Americans. Prevoicing is favored by pauses and vowels.

some utterance-initial realizations, as well as cases where the speaker paused before the target word, usually due to disfluency or emphatic speech.

For all three groups, a significant interaction was observed between PRECEDING ENVIRONMENT and STOP TYPE. For Polish Americans, this interaction is explained by differences in what conditions prevoicing for the two voiced stops (see Figures 7 and 8). Prevoicing exhibits selective conditioning for underlying /d/, as it is favored by preceding vowels and pauses, but not voiceless segments. By contrast, derived [d] favors prevoicing irrespective of the voicing of the preceding segment, which further confirms that negative VOT is associated with this stop type, rather than resulting from a voicing assimilation process, as in the case of underlying /d/. This different conditioning of prevoicing for underlying /d/ and derived [d] is specific to Polish Americans. Non-Polish New Yorkers also exhibit a significant interaction of PRECEDING ENVIRONMENT and STOP TYPE, but this effect is driven by longer VOTs for voiceless stops following vowels.

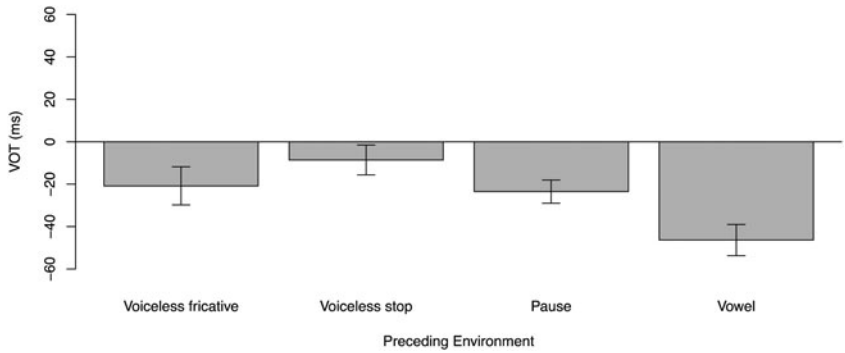


FIGURE 8. Mean VOT for derived [d] depending on preceding environment for second-generation Polish Americans. Prevoicing is favored regardless of context.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

VOT is also predicted by UTTERANCE POSITION. This effect is driven by significantly more positive VOTs in the utterance-initial position, relative to tokens that appeared utterance-medially or utterance-finally. A longer VOT indicates that underlying /d/ is realized as voiceless utterance-initially, which is consistent with reports for American English (Flege & Brown, 1982).

Lastly, GENDER was selected as a predictor of VOTs for second-generation Polish Americans only. Women tended to produce VOTs that were on average 5 msec longer than those produced by men. This gender difference is consistent with previous research, which has shown that English-speaking women produce longer VOTs than men do (Whiteside, Henry, & Dobbin, 2004), and Dutch-speaking men produce more frequent and longer prevoicing than women do (Van Alphen, 2007). This finding can be explained by men having larger vocal tracts, which allow more air to accumulate in prevoicing before transglottal pressure equalization inhibits voicing.

DISCUSSION

The results of this paper suggest that second-generation Polish Americans maintain separate VOT categories for English underlying /t d/ and derived [t d]. As is expected of English, an "aspiration" language, underlying stops contrast in aspiration. Derived stops, however, pattern in a way that is consistent with "voice" languages, such as Polish: derived [d] exhibits a strong preference for prevoicing, whereas derived [t] is unaspirated.

Crucially, the four-way separation of VOT categories is specific to second-generation Polish Americans. Non-Polish white New Yorkers maintain a two-way VOT contrast between long-lag underlying /t/, on the one hand, and short-lag underlying /d/ and derived [t d], on the other. Likewise, first-generation Polish Americans do not fully distinguish all stops, producing both voiced stops with prevoicing, but differentiating them from derived [t] and underlying /t/. This result points to the presence of a voicing bias in interference: underlying /d/ shows interference (in the form of prevoicing), whereas underlying /t/ is acquired successfully.

Remnants of this voicing bias can be observed in the second-generation's high proportion of prevoiced to short-lag VOTs. Prevoicing rates decrease across generations only for underlying /d/ (from 55% to 40%) and remain equally high for derived [d] (around 50%). These prevoicing rates are much higher than those observed for non-Polish white New Yorkers (i.e., 17%) and mirror those reported for other bilinguals whose L1 is a "voice" language. For example, Punjabi-English bilinguals employed prevoicing 50% of the time, compared to 5% to 20% for British English monolinguals (Heselwood & McChrystal, 2000). Similarly, MacKay et al. (2001) found four groups of Italian-English bilinguals to produce prevoicing for /b/ at rates ranging from 57% for early bilinguals to 79% for late bilinguals, compared to 31% for American English monolinguals.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

I argue that the voicing bias reflects differences in markedness relations between the bilinguals' L1 and L2. In Polish, voiced stops /b d g/ are believed to be marked with respect to their voiceless counterparts because of the presence of voicing throughout their articulation. In English, it is the aspirated stops /p t k/, characterized by a long VOT, that are thought to be marked with respect to the unaspirated series /b d g/. According to laryngeal realism (e.g., Honeybone, 2005; Iverson & Salmons, 1995; Kager et al., 2007), the marked member of a voicing pair carries a privative laryngeal specification, whereas its unmarked counterpart is thought to remain unspecified for any laryngeal features. These asymmetries in stops' featural specifications could account for the disproportionate pattern of L1 interference and transfer observed in this study: Polish /d/, specified for [voice], imposes a prevoiced realization on English /d/, which lacks this specification. By contrast, the absence of a laryngeal specification on Polish /t/ does not predict interference. The monolingual-like realization of English aspirated stops, therefore, relies on the acquisition of a new contrast, represented by the feature [spread glottis]. Within-language markedness distinctions can, therefore, account not only for the voicing bias toward interference with voiced stops, but also for the successful realization of the English long-lag VOT category by first-generation speakers. Note that the role of within-language markedness in predicting interference does not rest on any one phonological representation as divergent markedness relations could also presumably be deduced from the phonetic representation of L1 and L2 stop categories.

Within-language markedness thus emerges as a key determinant of interference. Previous work in language contact has focused on cross-linguistic markedness that translates into ease of learning (Thomason, 2010). TH-stopping represents this kind of cross-linguistic markedness, as seen in its generational decline in the speech of Polish Americans. Yet, a study of stopping rates alone, without any consideration of within-language markedness, would only provide a partial understanding of the interference effects observed here. Within-language markedness relations may also apply to cases of transfer reported elsewhere. For example, although cross-linguistic markedness explains the presence of TH-stopping for Louisiana Cajuns (Dubois & Horvath, 1998, 2000), it does not account for the maintenance of other French-influenced variables, such as nasalization on vowels, because lack of nasalization is not considered to be cross-linguistically marked. A linguistic motivation may be missing altogether, especially for speakers who are not bilingual, but this case of transfer could possibly be explained by differences in the phonological representation of (nasal) vowels in French and English. Although the exact outcome of interference will be ultimately subject to social considerations, within-language markedness may explain why some acoustic measures, such as VOT, are more prone to interference than others are.

Interference vs. transfer

Contrasting cross-generational profiles in the realization of VOTs suggest fundamental differences in the nature of Polish-like voicing for the two groups

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

of Polish Americans. For the first generation, Polish-like VOTs appear to be a product of L1 interference. These speakers manifest Polish-like voicing in underlying /d/ as well as in derived stops, which themselves result from L1 interference. Therefore, the first generation exhibits an involuntary, across-the-board implementation of Polish-like voicing, except for underlying /t/. Second-generation Polish Americans, on the other hand, relegate Polish-like voicing to derived stops. This selective presence of a non-native contrast suggests a degree of control that no longer resembles L1 interference. Rather, Polish-like VOTs in derived stops represent a case of transfer that is both linguistically and socially motivated.

At the linguistic level, transfer of Polish-like VOTs serves to preserve a contrast between underlying and derived stops. "Incomplete neutralization" is well studied and refers to the maintenance of slight phonetic differences between sounds that underlyingly represent different phonemes, but on the surface undergo a process that positionally obliterates the contrast, as in, for example, final devoicing (Warner, Jongman, Sereno, & Kemps, 2004). Similarly, the dichotomy between underlying and derived stops observed in this study may result from speakers' attempt to differentiate the two stop types as not corresponding to the same phoneme in their underlying representations. This interpretation is further motivated by Zhao's (2010) findings, which suggest that underlying /d/ and stop-like /ð/ differ on a number of characteristics, most likely resulting from differences in place or articulation between alveolar /d/ and dental stop-like /ð/. Thus, TH-stopping in general does not involve a one-to-one "substitution" of a stop for a fricative, as the common definition of the process might suggest. Instead, it produces a category of sounds that sustains some characteristics of the interdental fricative, thus only partially neutralizing the contrast between /θ ð/ and /t d/.

Although the transfer of Polish-like VOTs serves a linguistic role in the partial maintenance of a contrast, the progression from interference to transfer is simultaneously affected by social factors. Previous work has shown that TH-stopping may have gained a new meaning for Polish New Yorkers, as its rates correlate with speakers' ethnic orientation scores (Newlin-Łukowicz, 2013). Specifically, the most frequent stoppers lead a Polish lifestyle, engage in Polish cultural activities, and have Polish friends in New York City as well as in Poland. The least frequent stoppers, on the other hand, are removed from the Polish community and have predominantly non-Polish friendship circles. If TH-stopping carries some sort of ethnic connotations for Polish New Yorkers, then its phonetic implementation with Polish-like voicing, prevoicing for [d] and lack of aspiration on [t], further reinforces its (acoustic and social) salience. Therefore, even though the overall rates of TH-stopping decline across generations, the salience of derived stops is strengthened by the fact that underlying stops become more English-like. A similar reallocation of *t*-retroflexion to more salient contexts was observed for young second-generation Punjabi-English bilinguals in London (Sharma & Sankaran, 2011). Although the use of *t*-retroflexion decreased for the second generation compared to the first, retroflex /t/ gained a more "fortis" quality[5] and became restricted to the (more salient) word-initial position. Sharma

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

and Sankaran (2011) ascribed the focused retention of *t*-retroflexion to the second generation's orientation toward a local British Asian community, rather than the Indian motherland of their parents.

Similar cross-generational differences are observed for the speakers in this study and could explain the second generation's restriction of Polish-like voicing to derived contexts. The obvious distinction between the two generations is that of nativeness and immigrant status: first-generation speakers were born and raised in Poland and thus speak with an accent that identifies them as foreign-born. In most cases, they were forced to immigrate for economic reasons, leaving family and friends behind. Often not speaking a word of English and having no experience with diverse settings, they had to learn to navigate New York City's complex racial and ethnic relations. At some point or other, they came in contact with second-generation Polish Americans, and though some regarded their common background as basis for immediate friendship, others saw the "Americanized ways" of the second generation as alienating. The following excerpt, taken from a sociolinguistic interview with one of the speakers in this study, illustrates the disjunct between the second generation's "romantic" idea of Poland and their strong Polish identity, and the first generation's disenchantment with Poland, as described by a first-generation woman.

> **Magda**: When you're born here you have this very romantic idea of what Poland should be. You see these, Mazowsze, these people just. Just dancing beautifully, and *you see this, these beautiful forests and lakes and you think it's so romantic and so nice and lovely and you see these beautiful, Polish hairstyles and dresses and you think it's all great*, but you don't really know the dark side once you live there. So like it's. And it's really funny cause you don't wanna. cause it is true that that's, so there's a lot of beauty to it but it's funny that you just. *You don't wanna see the dark side* and you don't wanna know about it and you're not gonna even try, so I'm not gonna, talk to you about it and they, and they want, and *they feel like they're more Polish than I am*. In a way because they love it way more than I do, just simply because I know the, all the, all the sides how it fits together so it's always so funny to me how I meet these Polish people born here. Like "*Poland is the best, it's the best country to live in!*" *So why don't you go there?*

Second-generation Polish Americans were born in New York City and spoke English from age 5 onward. Many of them grew up in Polish neighborhoods, led Polish lifestyles, attended Polish supplementary schools and visited Poland every one to two years. Despite a heavily Polish upbringing, second-generation Polish Americans were exposed to racial diversity all their lives and many of them stressed being able to relate to other second-generation New Yorkers better than to first-generation Poles, who are often seen as unassimilated. This view is illustrated in the following interview excerpt, which captures a second-generation woman's idea of first-generation Poles.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

> **Alicja**: *They like to think they're better*. Especially the ones that came here, when they were older, and *they feel this connection, this deeper connection with their, with their motherland*. I don't. I feel more, actually I have more loyalty, although loyalty's maybe not the best word. *Solidarity towards America*. Actually. *I relate to the greater culture much more*, even though I never had an American friend. Much more than to Polish people. I find them to be *really narrow, more narrow minded, their view of the world is more limited*. In America, it's more open.

These excerpts provide only a small window into the cross-generational differences among Polish Americans and should not be taken to represent the views of every member of the community. Although this evidence is qualitative in nature, these cultural differences across generations could conceivably account for the linguistic differences between interference and transfer observed here. Specifically, the second generation's restriction of Polish-like voicing to derived contexts linguistically differentiates them from the first generation. Although both generations manifest Polish-influenced features in their speech, the accent of the first generation can be considered more *global*, as it is unrestricted, whereas the accent of the second generation is more *focused* and disproportionately applies to derived contexts. These social and linguistic differences across generations further suggest that TH-stopping, realized with Polish-like voicing, may act as an ethnic identity marker for the second generation only.

Models of interference

The results from this study suggest that L1 interference and transfer are variable and asymmetrical, with interference favoring voiced over voiceless stops, and transfer preferring derived over underlying stops. The observed voicing bias is consistent with existing theories of L2 acquisition, as it reflects within-language markedness. However, the *variable* presence of negative and short-lag VOTs for English voiced stops is not predicted by existing theories, which assume that once established, L2 phonetic categories are used across the board. This deterministic view is reflected in the fact that previous work on L2 interference almost exclusively compared bilinguals' cross-linguistic productions, assuming that within-language categories were stable. In addition, the dominant models of interference, such as PAM nor SLM, are not meant to account for cases of transfer as they do not incorporate social information. However, because features that become "transferred" commonly derive from interference patterns, a unified model of interference and transfer would be desirable. In fact, PAM and SLM could be extended to account for variation if they assumed a more nuanced understanding of L1/L2 phonetic categories and the grammar that governs their choice.

This paper motivates a view of bilingual phonology as governed by a variable grammar with access to phonological and social information. The variable realization of VOTs, along with the reported separation of VOT categories for underlying and derived stops, demonstrate that L1 phonetic categories can affect the production of "similar" L2 categories, thus providing ground for transfer. Crucially, this transfer is socially driven: derived stops with Polish-like voicing

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

likely serve as a type of ethnic marker for second-generation Polish Americans (Newlin-Łukowicz, 2013). If we assume that the distribution of L1/L2 phonetic categories is subject to variable rules or constraints of the sort discussed by Labov (1969), Cedergren and Sankoff (1974), Sankoff (1978), and Guy (1991), we can achieve a unified account of the transition from the global effects of interference to the focused presence of transfer. In fact, the transfer effect observed for second-generation Polish Americans falls out of the interference pattern reported for the first generation in that it exhibits remnants of the voicing bias. The preference for transfer in derived stops, on the other hand, reflects the second generation's successful acquisition of English-like VOTs in underlying stops, largely freed from interference, as well as the development of a possible social meaning attached to TH-stopping in the Polish community in New York City. Transfer effects, previously understood as decreased rates of interference (Thomason & Kaufman, 1988; Van Coetsem, 1988), would then emerge as a reweighting of the constraints that trigger interference for the first generation, modulated by phonological and social factors.

NOTES

1. Empirical evidence for the presence of the feature [voice] in some languages, and [spread glottis] in others, comes from the study of boundary phenomena in "voice" and "aspiration" languages (Iverson & Salmons, 1995), assimilation biases (Iverson & Salmons, 1999), language change (Honeybone, 2005), and language acquisition (Kager et al., 2007).

2. Eleven of the generation one speakers came to the United States at ages 10 to 15, and the remaining four arrived at age 8. Although the latter speakers could be classified as generation 1.5, they pattern closely with the other generation one speakers in terms of TH-stopping rates. They also express identification with other first generationers in sociolinguistic interviews because of their mutual immigrant experience.

3. The short stories overall included many more function words but VOT was only measured for the ones designed to be tested.

4. The deviations between the targeted and produced stop tokens are concentrated in particular positions, with larger deviations in the final position. This suggests that in the final position (1) TH-stopping is less likely (for derived stops) and (2) unreleased realizations are more likely (for underlying stops). This result is expected, as unreleased final stops are common in American English.

5. Sharma and Sankaran (2011) do not elaborate on their observation about the acoustic quality of retroflex /t/, nor do they speculate about the acoustic correlates that correspond to the described "fortis" quality.

REFERENCES

Adamson, Hugh Douglas, & Regan, Vera M. (1991). The acquisition of community speech norms by Asian immigrants learning English as a second language: A preliminary study. *Studies in Second Language Acquisition* 13:1–22.

Baayen, R. Harald. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.

Babbitt, Eugene Howard. (1896). The language of the lower classes in New York City and vicinity. *Dialect Notes* 1:457–464.

Bates, Douglas, & Maechler, Martin. (2009). lme4: Linear mixed-effects models using S4 classes. Available at: http://CRAN.R-project.org/package=lme4. Accessed February 6, 2014.

Bayley, Robert, & Preston, Dennis. (1996). *Second language acquisition and linguistic variation*. Amsterdam: John Benjamins.

Becker, Kara. (2010). *Social conflict and social practice on the Lower East Side: A study of regional dialect features in New York City English*. Ph.D. dissertation, New York University.

Best, Catherine. (1995). A direct realist view of cross-language speech perception: New directions in research and theory. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues*. Baltimore: York Press. 171–204.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

Best, Catherine, & Tyler, Michael. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro & O.-S. Bohn (eds.), *Second language speech learning: The role of language experience in speech perception and production*. Amsterdam: John Benjamins. 13–34.
Boersma, Paul, & Weenink, David. (2014). Praat: Doing phonetics by computer. Version 5.3.63. Available at: http://www.praat.org/. Accessed January 24, 2014.
Britain, David. (2002). Diffusion, leveling, simplification and reallocation in past tense BE in the English Fens. *Journal of Sociolinguistics* 6(1):16–43.
Cedergren, Henrietta I., & Sankoff, David. (1974). Variable rules: Performance as a statistical reflection of competence. *Language* 50:333–355.
Chang, Charles, Yao, Yao, Haynes, Erin F., & Rhodes, Russell. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *Journal of the Acoustical Society of America* 129:3964–3980.
Docherty, Gerard. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris Publications.
Drummond, Rob. (2012). Aspects of identity in a second language: ING variation in the speech of Polish migrants living in Manchester, UK. *Language Variation and Change* 24:107–133.
Dubois, Sylvie, & Horvath, Barbara. (1998). Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change* 10:245–261.
_____. (2000). When the music changes, you change too: Gender and language change in Cajun English. *Language Variation and Change* 11:287–313.
Flege, James Emil. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press. 233–272.
Flege, James Emil, & Brown, William Samuel. (1982). The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance. *Journal of Phonetics* 10:335–345.
Flege, James Emil, & Eefting, Wieke. (1987). The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics* 15:67–83.
Guy, Gregory. (1991). Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3:223–239.
Hall-Lew, Lauren. (2009). *Ethnicity and phonetic variation in a San Francisco neighborhood.* Ph.D. dissertation, Stanford University.
Heselwood, Barry, & McChrystal, Louise. (2000). Gender, accent featured and voicing in Panjabi-English bilingual children. *Leeds Working Papers in Linguistics and Phonetics* 8:45–70.
Honeybone, Patrick. (2005). Diachronic evidence in segmental phonology: The case of laryngeal specifications. In M. van Oostendorp & J. van de Weijer (eds.), *The internal organization of phonological segments*. Berlin: Mouton de Gruyter. 319–354.
Iverson, Gregory K., & Salmons, Joseph C. (1995). aspiration and laryngeal representation in Germanic. *Phonology* 12:369–396.
_____. (1999). Glottal spreading bias in Germanic. *Linguistische Berichte* 178:135–151.
Jacewicz, Ewa, Fox, Robert Allen, & Lyle, Samantha. (2009). Variation in stop consonant voicing in two regional varieties of American English. *Journal of the International Phonetic Association* 39:313–334.
Kager, René, van der Feest, Suzanne, Fikkert, Paula, Kerkhoff, Annemarie, & Zamuner, Tania. (2007). Representations of [voice]: Evidence from acquisition. In J. M. van de Weijer & E. J. van der Torre (eds.), *Voicing in Dutch*. Amsterdam: John Benjamins. 41–80.
Keating, Patricia Ann. (1980). *A phonetic study of a voicing contrast in Polish.* Ph.D. dissertation, Brown University.
Keating, Patricia Ann, Mikoś, Micheal J., & Ganong, William F. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America* 70:1261–1271.
Labov, William. (1966). *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
_____. (1969). Contraction, deletion, and inherent variability of the English copula. *Language* 45:715–762.
Lisker, Leigh, & Abramson, Arthur S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20:384–422.
Lombardi, Linda. (2003). Second language data and constraints on Manner: Explaining substitutions for the English interdentals. *Second Language Research* 19:225–250.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

MacKay, Ian R. A., Flege, James Emil, Piske, Thorsten, & Carlo Schirru. (2001). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America* 110:516–528.

Mendoza-Denton, Norma. (2002). Language and identity. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Malden: Blackwell Publishers. 475–499.

Nagy, Naomi, & Kochetov, Alexei. (2013). Voice onset time across the generations: A cross-linguistic study of contact-induced change. In P. Siemund, I. Gogolin, M. Schulz, & J. Davydova (eds.), *Multilingualism and language contact in urban areas: Acquisition—development—teaching—communication*. Amsterdam: John Benjamins. 19–38.

Newlin-Łukowicz, Luiza. (2013). TH-stopping in New York City: Substrate effect turned ethnic marker? *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV 41* 19(2):151–160.

Preston, Dennis. (1989). *Sociolinguistics and second language acquisition*. Oxford: Basil Blackwell.

R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria. Available at: http://www.R-project.org. Accessed February 6, 2014.

Sankoff, David. (1978). Probability and linguistic variation. *Synthèse* 37:217–238.

Sankoff, Gillian. (2002). Linguistic outcomes of language contact. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 638–668.

Schleef, Erik, Meyerhoff, Miriam, & Clark, Lynn. (2011). Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London. *English World-wide* 32(2):206–236.

Sharma, Devyani. (2005). Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition* 27:535–566.

Sharma, Devyani, & Sankaran, Lavanya. (2011). Cognitive and social forces in dialect shift: Gradual change in London Asian speech. *Language Variation and Change* 23:399–428.

Simon, Ellen. (2010). *Voicing in contrast. Acquiring a second language laryngeal system*. Gent: Academia Press.

Sundara, Megha. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *Journal of the Acoustical Society of America* 118:1026–1037.

Thomason, Sarah Grey. (2003). Contact as a source of language change. In B. D. Joseph & R. D. Janda (eds.), *The handbook of historical linguistics*. Malden: Wiley Blackwell. 687–712.

_____. (2010). Contact explanations in linguistics. In R. Hickey (ed.), *The handbook of language contact*. Malden: Wiley Blackwell. 31–47.

Thomason, Sarah Grey, & Kaufman, Terrence. (1988). *Language contact, creolization, and genetic linguistics*. Berkley: University of California Press.

Van Alphen, P. M. (2007). Prevoicing in Dutch initial plosives: Production, perception, and word recognition. In J. Van de Weijer & E. J. van der Torre (eds.), *Voicing in Dutch. (De)voicing—phonology, phonetics, and psycholinguistics*. Amsterdam: Benjamins. 99–124.

Van Coetsem, Frans. (1988). *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

Warner, Natasha, Jongman, Allard, Sereno, Joan, & Kemps, Rachel. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics* 32:251–276.

Wells, John C. (1982). *Accents of English*. New York: Cambridge University Press.

Whiteside, Sandra P., Henry, Luisa, & Dobbin, Rachel. (2004). Sex differences in voice onset time: A developmental study of phonetic context effects in British English. *Journal of the Acoustical Society of America* 116:1179–1183.

Wolfram, Walt, Carter, Phillip, & Moriello, Beckie. (2004). Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8:339–358.

Wolfram, Walt, & Hatfield, Deborah. (1984). *Tense marking in second language learning: Patterns of spoken and written English in a Vietnamese community*. Washington, DC: Center for Applied Linguistics.

Zhao, Sherry Y. (2010). Stop-like modification of the dental fricative /ð/: An acoustic analysis. *Journal of the Acoustical Society of America* 128:2009–2020.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

APPENDIX

*Short stories (tokens in **bold**)*

1. **Talk** about a funny story! When I was still new in the business, a guy offered me a gig at his comedy club in Santa Monica. **Thrilled** to finally have a chance to prove myself, I agreed to do live stand-up **there**.
   I get to the address and I immediately notice that I'm not in a good part of town. Not sure if I feel safe. **Dark**, gloomy buildings surround the club. But hey, at least I don't have to worry about parking. **There's** plenty of space by the Ballona Creek bank.
   I spot the illuminated silhouettes of women flashing on the roof of the club. **That**, coupled with a neon sign above the **door** makes me a bit suspicious.
   "Dance till you drop," **that's** what the sign says. I didn't think **there** would be any dancing at a jazz club, but whatever.
   Then I notice a man out front handing out religious pamphlets.
   "**This** is not the house of **God**!," he yells. "**Think** twice before you walk in. Every time you walk through that door a part of your soul dies. Piece by piece. In the end, who do you want the Lord to save?"
   Deep inside of me, I start feeling uneasy. Not paying attention to **this** weirdo, I march to the **door. There**, I'm stopped in my tracks by the bouncer. I tell him that I'm there to perform.
   "Are you here for ladies' night?" he asks. I had no idea this was an all-female event. I think to myself: "**This** place really supports female comics."
   The bouncer motions into a cellar. I go into a rather small room and meet the other comics **there**. My first thought: Why are they all so attractive? I wonder if they'll be telling jokes about being single and dating like me. My second thought: Why are they only wearing bras and underwear?
   Suddenly, the feeling comes over me **that** I had once before when I was applying lip liner in a poorly lit bathroom at a TGI Friday's and a man emerged from the stall — I'm in the wrong place!
   They all think I'm a stripper! The bouncer, the waitress, and the audience **think** I'm here to strip! How did I not see **that?**
   Of course, I was flattered. Who wouldn't be? And when I found out the prize was six **thousand** dollars, I considered entering. Then, I remembered the high-waisted panties I was wearing and decided to stick with comedy. **Days** like these should either be forgotten, or laughed at.
2. It was my first time on the Late Show with **David** Letterman, and I was doing my best to control my nerves. I was so excited to be **there**! I knew the crowd was expecting a cheap **thrill**: **Tom** Cruise was scheduled to talk with **David**.
   Thousands of fans were lined up in front of the CBS building.
   That night Tom Cruise was pumped. Almost as pumped as he was during the Oprah Winfrey show. He had become the butt of all jokes since **this** performance, having jumped off a sofa and onto an antique **table**. And all of this just to proclaim his love for Katie Holmes, his better half.
   So here I am, doing my best to stay cool before going on. All of a sudden, someone runs offstage and makes a beeline for me. Of course it's **Tom**. He's sweating and breathing heavily like he has just won a prizefight. He grabs my hand and looks **deep** into my eyes, like **this**. I notice a nervous twitch. Or was that a wink?

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

"Whoo!" he yells. "Whoo?" I try to mimic but my voice does not sound that **deep**. I look down. I see that Tom's arm left a sweat mark on my new **dark** leather jacket. (Don't even ask about the price). "Dingy is the new cool," I console myself. "**Take** it off!" – I hear the stylist's voice hollering behind me. "Drop **that**!" – she repeats, pointing at the jacket, and vanishes behind a puff of make-up that now surrounds my face. I *so* do not want to do **this**. I'm convinced that this chic **three** thousand dollar jacket is my ticket to stardom.

I can feel Tom grip my hand again. "I promise **that** you'll do great. **Tell** me, are you nervous?" he asks. "A little bit. **That's** the only jacket I got."

Tom goes, "Do I need to have a serious **talk** with you? It's not about the jacket. And it's not about anything else you may or may not have . . . You have your own fears to beat! That's all it **takes.**" I thank **Tom** for his words of wisdom. He yells an excited "You bet!" before he disappears. I didn't even see him leave.

He says I'll do great but how can he **tell**? And how can he *promise* **that**?

I am now about to perform in front of a crowd of people who apparently filled Tom Cruise with pure rocket fuel, and I'm freaking out even more! I **tell** myself that I got nothing to prove. **David** Letterman—who I totally forgot about—is getting ready to introduce me.

I don't remember much of what happened on the show. But I discovered one **thing**. The wack **thing** about being a comedian is you don't really have to be **there** as long as the jokes show up, which, luckily, happened on this landmark **day**. And for that realization, I owe Tom Cruise a generous ***thank*** *you*.

3. **Thing** is, I got into show business for the parties. When I joined the cast of MADtv in the year two **thousand**, I was in my party phase.

One historic **day** I posted signs at all of the show's offices in the production wing: "Party at My House! Bring Anyone!" It would prove to be poorly worded.

At 8:30 on the dot, the doorbell rang.

Doors open and my first guest walks in. She's a demure-looking stranger in her 60 s, I **think**. Weirdly **thrilled** to meet me. She introduces herself as Deb.

Deb goes, "I'm Tina's **niece.**"

"**Tina** . . . ," I stutter.

"She works **there,** at MAD."

I pretend to know who Tina is. We go into the kitchen, and I make **Deb** a drink.

"One, two . . ." (gulp) "**three**!" She **takes** a sip and immediately spits it out into the sink.

"Bad?" I ask. "I wouldn't know. I've never had alcohol in my life! But hey, let's kick **this** party off with a bang!"

Three more strangers arrive. They all come in saying: "we're friends with **Tina.**" I'm starting to get concerned. Who is this mysterious **Tina** and how many people has she invited? I hate being kept in the **dark**.

I'm right to be worried because by 10 p.m. the air is heavy with social ineptitude. **There** are 25 of my friends, 30 friends of Tina's and no Tina. Time is passing slowly as I'm trying to keep the party alive. I try to get people to sing. But a few have already passed out on the sofa bed.

Others are playing a drinking game with **Deb**. They're sitting on their knees. Throwing dice.

I see two guys challenge each other to thread a string of peas. Just like you would a bead. In the meantime, their buddy is placing a bid.

"What's the prize?" I ask. "The winner gets ***this***!," the guy answers, pointing at my

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press

vintage baseball bat!

Eventually, I learned that Tina and her friends were enrolled in an institute called Top Rank, where they were taking classes in confidence building. Their homework was to attend a party and find someone to **talk** to. Apparently, Tina, who holds an administrative position at MADtv, passed along my invitation—to all 100 students. And here's proof!

Around 11, in walks a big, smiling face. She yells, "Hi, I'm Tina! I'm the one you have to **thank** for getting your party started!" She then hands me a picnic **table**. "There! A housewarming gift! I always like to do nice **things**."

A . . . **table**?! **This** is when I knew she was a crazy person. "**Thank** you?," I hesitated.

At the end of the night, I gave Tina and each of her friends a class evaluation: everyone got an F in networking, except Deb, who got credit for being punctual.

As for myself, I decided it was time to turn over a new leaf. No more parties with unexpected guests.

https://doi.org/10.1017/S0954394514000167 Published online by Cambridge University Press