When two language learners converse in a lingua franca, their phonetic production is often affected by transfer of phonological categories from their native languages; this results in distinct, predictable difficulties in pronunciation for each speaker (Flege et al. 2003). Such a context provides a rich environment for investigating the degree to which speakers can tune their pronunciation to one another in discourse (phonetic alignment). Trofimovich and Kennedy (2014) found that speakers from a variety of language backgrounds using English as a lingua franca sounded more similar to one another after conversing, suggesting that the speakers’ pronunciation does in some respect align. In this study, we ask whether those categories transferred from the native language of each conversant can be aligned, and we examine alignment in real time rather than post hoc. Since Costa and colleagues (2008:550) hypothesize that typological distance of the conversants’ native languages would inversely affect alignment, we investigate phonetic alignment among speakers with controlled non-native language backgrounds. The current study uses a new corpus of English as a lingua franca to provide a fine-grained analysis of phonetic alignment to two key vowel contrasts as a function of time and speech style.

The Nijmegen Corpus of Spanish English (NCSE; Kouwenhoven et al. Forthcoming) consists of recorded conversations between 34 Spaniards and 2 Dutch confederates in English across both formal and informal speech styles. The manipulation of speech style permits an analysis of phonetic alignment in two discourse contexts with distinct situational stressors. Additionally, Spaniards’ English proficiencies were rated by an experienced TESOL teacher using the Common European Framework of Reference for languages, permitting us to evaluate alignment as a function of proficiency as well.

We tracked Spanish speakers’ production over the course of both formal and informal conversation in NCSE, focusing on two key English vowel contrasts. One contrast, /ɪ/ vs. /i/, was chosen because it is reportedly difficult for Spanish learners of English but simple for Dutch learners of English (Boomershine 2013; Levey and Cruz 2004). The second, /ɛ/ vs. /æ/, was chosen for the opposite reason. While Spanish does not have either /ɛ/ or /æ/, these tokens are pulled toward the space of Spanish tense vowels /e/ and /a/, nonetheless creating a reliable contrast between the two. Dutch speakers of English, however, merge these two categories (Wang and van Huneven 2006): tokens are pulled toward the Dutch vowel /ɛ/, which occupies the lower front quadrant of the vowel space in their L1 (cf. Booij 1995:5).

To examine the effect of exposure on production, we rescaled time and created ten equally-spaced bins by speaker and speech style. We extracted all stressed tokens of the four vowels of interest (N_/ɪ_/=8,482; N_/i_/=18,627; N_/ɛ_/=20,506; N_/æ_/=17,343) and grouped them by contrast. Next, we predicted logmean-normalized F1 and F2 frequencies by vowel with separate MANOVAs for each time bin and condition, excluding time bins with fewer than 20 tokens. Pillai scores from these MANOVAs were then used to approximate vowel contrast (0 meaning no contrast, and higher values meaning more contrast; cf. Hay et al. 2006; Nycz and Hall-Lew 2013).

First, we note an effect of L1 category transfer: Spaniards on average do not distinguish /ɪ/ from /i/ (x̄=0.15, s=0.11), but do make a distinction between /ɛ/ and /æ/ (x̄pilla=0.47, spilla=0.17). Second, we compare the two contrasts as a function of time, style, and proficiency. Pillai scores were entered as the dependent variable for linear mixed effects models (Bates et al. 2015); a separate model was run for each vowel contrast (N_/ɪ_/=658; N_/i_/=660). These models revealed two key findings. First, Spanish learners of English aligned to their Dutch interlocutors primarily in informal speech. While the /ɪ/-/i/ contrast is more merged in informal speech overall (β = -0.08, t = -4.04), it separates gradually over time in this context (β = 0.01, t = 3.32; see Figure 1). Speakers increase the degree of merger for the /ɛ/-/æ/ categories when speaking informally (β = -0.15, t = -11.74), but this does not change significantly over time. Second, alignment is modulated by proficiency in the lingua franca, particularly for the /ɛ/-/æ/ contrast. In the latter, degree of merger increases with proficiency in English (β = -0.17, t = -3.74; see Figure 2), implying that high-proficiency speakers suspend a phonological category distinction they are capable of making in order to align to the production patterns in the discourse context.

To summarize, non-native speakers from distinct L1s attune to their interlocutors, even in cases where doing so causes their L2 speech to deviate from more native-like production. Alignment is, however, subject to speech style and proficiency, being more prevalent in informal speech and increasing with higher L2 proficiency. Given that Spanish and Dutch are typologically distinct and have
very different phonologies, our findings cast doubt on the hypothesis of Costa and colleagues (2008), and instead suggest that the mediating factor in non-native discourse alignment is a combination of situational stressors and one’s experience with and proficiency in the lingua franca. We interpret these findings in a cognitive processing framework where top-down, discourse context-sensitive information interacts with bottom-up processes of discourse alignment to produce the observed behavior.

**Figure 1:**
Pillai scores by Speech Style, Time, and Contrast

**Figure 2:**
Pillai scores by Proficiency and Contrast

References: