Proceedings of the Virtual Transnational Linguistic Project on

Speech Production and Perception and its Cognitive Processing

2023

Effects of Social Provenance in Vowel Accommodation in Parisian Varieties

Inge Akyaa[1], Olivia Caffin[2], Suyuan Dong[2], Léna Gaubert[2]

[1]*University of Cologne, Germany*
[2]*University of Sorbonne Nouvelle, France*

iakyaa@gmail.uni-koeln.de, olivia.caffin@sorbonne-nouvelle.fr,
suyuan.dong@sorbonne-nouvelle.fr, lena.gaubert@sorbonne-nouvelle.fr

Abstract

This study explores the influence of social factors on vowel accommodation in Parisian varieties of French. Eight female native French speakers from Paris and Saint-Denis engaged in a map task, adopting roles as instructors and followers, and producing target vowels /a/, /i/, and /u/. Findings suggest significant vowel convergence among speakers with similar backgrounds, while speakers from different backgrounds displayed a combination of convergence and divergence. These results suggest that speech accommodation is a multifaceted process, reflecting both an automatic process and strategic behavior, deeply influenced by social factors.

Keywords: Vocalic Accommodation, Socio-Phonetics, Conversational Interaction, Map Task, Multicultural Context

1. Introduction

Speech accommodation is a complex and dynamic process, conditioned by the cognitive organization of a larger language system. This phenomenon, characterized by speakers' unconscious adaptation of their acoustic-phonetic attributes to those of their interlocutor, has been extensively documented in previous studies (Babel & Bulatov, 2012; Pardo, 2022). Several approaches are used to describe the multiple directions and extent of this phenomenon, such as alignment, convergence, divergence, imitation, and mimicry. Despite the terminological differences, these approaches underscore the nature of "inter/cognitive and/or external/social mechanisms" driving changes during the communicative procedure (Pardo, 2022). Besides, this phenomenon is manifested through various acoustic characteristics, such as VOT, f0, and vowel imitation. It is also suggested that this phenomenon is not only phonetic but also deeply influenced by other task-specific and talker-specific factors, such as gender, age, personality, linguistic background, cultural identity, and role during interaction. Furthermore, the extent and nature of accommodation are subject to both short/long term (*idem.*).

The present study aims to delve into the social dimensions of speech accommodation, particularly examining the influence of cultural background among speakers of different varieties within the same language, and also the impact of interactional roles during the task on the accommodation process.

1.1. Social factors

Social factors that influence vocal accommodation include social status, power dynamics, social norms, visual cues, familiarity, and social identity. Furthermore, social knowledge about the model talkers can inhibit or facilitate the accommodation process (Babel, 2009a, 2009b; Babel & Bulatov, 2012; Kim et al., 2012). For example, individuals may unconsciously mimic the speech patterns of someone they perceive as having higher social status or power (Ostrand & Chordoff, 2021). In the context of North and South Korean immigration in London, Evans (2023) states that North Korean refugees accommodate closely to South Korean immigrants, especially if they left North Korea throughout their childhood. However, because of diverging identities, both communities tend to diverge during conversational speech.

Immigration in the Parisian suburbs has created a strong multicultural and multilingual context: the region of Île-de-France has the highest proportion of individuals having a "bilingual daily life", with one child out of four being in contact with another language than French at home (Gadet 2007: 128). While studying the dialect accommodation of Cherokee groups in Western North Carolina, Anderson (1997) showed that the less isolated group (and therefore the group that was the most in contact with English-speaking and white communities) showed stronger accommodation, specifically with the monophthongal vowel [a:]. This observation necessitates a focus on the dynamics inherent within the multilingual context under consideration.

1.2. Typological background

Suburban Parisian French, while prevalent in various suburbs, is also spoken in some intra-muros parts of Paris. For this study, the focus will be on speakers from Saint-Denis, the largest city in the suburbs in the northeast of Paris, known for its multiculturalism. This variety of French has been the subject of numerous studies aiming to describe its phonetic characteristics (Paternostro, 2012; Candea, 2016). A notable feature of this language, derived in part from what is known as "popular French" of the working class, includes consonant variations such as simplification of consonantal clusters, consonantal assimilation, and palatalization. The vowels, on the other hand, are distinctively influenced by a more multicultural context, particularly by Maghrebian languages. This influence is evident in the shortening and occasional elision of vowels, leading to vowel reduction, which is a phenomenon relatively rare in French (Fagyal, 2010: 96). Gadet et al. (2017) also identify vocalic harmony and closed realization of vowels that are supposed to be open (e.g. <mère> /mɛʁ/ becomes /meʁ/). Despite the prevalent focus on consonants and prosody in existing research, there is a gap in studies addressing vowel variations in this variety. Considering this lacuna and the fact that the cardinal vowels /a/, /i/, and /u/ are not only present in both target varieties but also embody the multicultural attributes of the Maghrebian community, our study specifically focuses on the reduction of these vowels.

1.3. Hypothesis and predictions

1 – Interlocutors with the same socio-geographical background will display an increased likelihood of observing vowel convergence.

2 – Interlocutors originating from different socio-geographical backgrounds may display different accommodation behaviors according to their subjective representations of their personal and interlocutor identity, of the context, and the perceived linguistic power dynamics within the interaction.
3 – Participants may show different accommodative behaviors depending on their role as *Follower* or as *Instructor*.

2. Methods

The experiment is designed to test four conditions: Paris-Paris (PP), Suburb-suburb (SS), Paris-Suburb (PS) and Suburb-Paris (SP). Participants assumed two roles in the map task: Instructor and Follower.

2.1. Participants

To minimize the gender effects in speech, this study engaged eight female native French speakers (aged 16 to 18): four were born and raised in Paris, while the other four are of Maghrebian origins from Saint-Denis. All participants were high school students with no specialized linguistic training and exhibited normal speech and hearing capabilities.

2.2. Procedure

The experiment was conducted in a sound-insulated room. Participants were seated facing one another, equipped with high-quality microphones and a computer monitor on the desk to view instructions. The passing order was randomized yet structured to ensure equal representation of each condition, following the sequence: s-p-p-s-s-p-p-s. Each participant enrolled two times by playing both *Instructor* and *Follower* in turn (1-2, 2-3, …, 8-1).
Upon arrival, participants engaged in an initial ice-breaking session involving self-introductions about names, ages, study programs, and residential areas. Furthermore, participants were prompted to share three popular expressions from their schools. This exchange aimed to build connections, set a conversational tone among them, and activate the use of their daily language in preparing for the subsequent tasks.
Before starting the main task, participants engaged in a preliminary testing exercise to ensure their comprehension of the procedure. Then during the pre-task session, the dyad in role needed to read individually a staging text written in the form of a diary, containing words with target vowels, appearing on the computer screen. This activity aimed to immerse participants in the map-task scenario, and the recordings serve also as a baseline for the vowel comparison.
In the main map task, participants were presented individually with a map on their computer: the instructor's map detailed with place names, while another with only road names and traffic information. The objective was to identify locations and purchase items needed collaboratively within one hour. Following the map task, each participant read another text, similar in structure to the pre-task, to reflect on the completed activity and repeat the target words. This served to assess any phonetic shifts post-interaction.
Participants completed a questionnaire regarding their feelings and perceptions of language variations during the whole experiment.

3. Possible results and discussion

The acoustic analysis of the pre-task productions of the cardinal vowels /a/, /i/, and /u/ aligns with the predicted vowel distribution. This result suggests that vowels in the Maghrebian variety of French exhibit a more centralized vocalic space compared to those in Parisian French.
When it comes to the map task, results may indicate a significant level of vowel convergence among participants sharing similar socio-geographical backgrounds (P-P and S-S dyads). In contrast, dyads of P-S and S-P participants displayed a mixture of both convergence and divergence. This phenomenon was observed in both directions (P → S or S → P). Questionnaire responses from these groups indicated a higher awareness of language identity, which appears to correlate with their divergence during the communicative interactions. However, the variation in roles (*Instructor* vs. *Follower*) within the map task may not provide significant differences in terms of vocalic accommodation. This suggests that individual representation and personality may exert a more substantial influence on speech patterns than the assigned roles within the interaction. Besides, the observation of post-task accommodation mirrors the previous studies.
Further findings may emerge beyond the primary focus. These include potential shifts in prosody, syntax, variations in lexical choice, and the use of supralinguistic features such as clicks to express attitude. These aspects, while not central to the current study, may offer intriguing directions for future research on speech accommodation in multilingual contexts.

4. References

Babel, Molly (2009a): Phonetic and Social Selectivity in Speech Accommodation.

Babel, Molly (2009b): Selective Vowel Imitation in Spontaneous Phonetic Accommodation. UC Berkeley PhonLab Annual Report 5(5) doi:10.5070/P76c95w485.

Babel, Molly & Dasha Bulatov (2012): The Role of Fundamental Frequency in Phonetic Accommodation. Language and Speech. SAGE Publications Ltd. 55(2). 231–248.

Candea, M. (2016). L'accent dit 'de banlieue', une mode? Étude auprès de lycéens en ZEP inscrits dans une dynamique de réussite scolaire.

Fagyal, Z. (2010). Rhythm types and the speech of working-class youth in a banlieue of Paris: the role of vowel elision and devoicing. Preston & Niedzielski (éds.), 91-132.

Gadet, F., Cappeau, P., Conein, B., Guerin, E., Moreno, A., Paternostro, R., ... & Wachs, S. (2017). Les parlers jeunes dans l'Ile-de-France multiculturelle. Éditions Ophrys.

Kim, Midam, William S. Horton & Ann R. Bradlow (2011): Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. Laboratory Phonology 2(1) doi:10.1515/labphon.2011.004.

Ostrand, Rachel & Eleanor Chodroff (2021): It's alignment all the way down, but not all the way up: Speakers align on some features but not others within a dialogue. Journal of Phonetics 88. 101074. doi:10.1016/j.wocn.2021.101074.

Paternostro, R. (2012). Aspects phonétiques du "français parisien multiculturel": innovation, créativité, métissage (s). Cahiers de l'AFLS, 17(2), 32-54.

Pardo, Jennifer S., Elisa Pellegrino, Volker Dellwo & Bernd Möbius (2022): Special issue: Vocal accommodation in speech communication. Journal of Phonetics 95. 101196. doi:10.1016/j.wocn.2022.101196.

Anderson, Bridget L. (1997): Adaptive sociophonetic strategies and dialect accommodation: /ay/ monophthongization in Cherokee English.

Evans, Bronwen G., Rakgyun Shin & Justin Lo (2023): Being North Korean in Koreatown: the role of identity & experience in accommodation in the London North Korean community.

Cross-Linguistic Imitation: English - Turkish

Erik[1], Dilara[1], Doğa[2], Nadya[2]
[1]Université Paris Cité
[2]Universität zu Köln

Introduction

Phonetic imitation (accommodation, convergence) can loosely be defined as "an increase in the perceptual or acoustic similarity between one speaker and another". This phenomenon has been studied extensively within speakers of a single language, and speakers have been shown to imitate consistently (Wade et al., 2021), but only a few have investigated cross-linguistic imitation.

Coles-Harris (2017) discusses some of the literature on imitation and assumptions on its nature, which found the basis of this study as well. They discuss the two primary views on the motivation for imitation, whether it is a wholly automatic process or one that is motivated by social factors. We, like Coles-Harris, view imitation as a largely automatic process that can be mediated by external, social factors (Jiang & Kennison, 2022; Zając & Rojczyk, 2014). This automatic process is based on a direct link between perception and production and it is this that suggests imitation would cross language boundaries.

Experiment

This study aims to explore the cross-linguistic persistence of imitation and whether a participant who imitates highly in their L1 also does so in an L2. We can then question whether this persistence would confer an advantage in language acquisition. This is reflected in our two research questions:

(1) Is the degree of imitation a person exhibits consistent across languages?

(2) Do "high imitators" have an advantage in phonetic and phonological second language acquisition?

Languages

The Turkish vowel system consists of eight vowels, six of which come in rounded-unrounded pairs (Sabev, 2019). Rounding is actively used to distinguish between vowel pairs of the same height and backness, and it is important in Turkish vowel harmony, making this feature key to proficient communication. English, on the other hand, makes greater use of position in the vowel space and duration to distinguish vowels. Rounding is present in English, though largely made redundant by the aforementioned cues, only in the differentiation between high and non-high front versus back vowels (de Jonge et al., 2022). It is for

this reason that Turkish and English are the chosen languages to examine cross-linguistic imitation in our study.

English, in our case Standard Southern British English (SSBE), has ten vowels that make little use of rounding contrastively. Rather, English vowels are primarily distinguished by duration and position. This combination provides us with two vowel systems that share some vowels but use them in different ways, as well as vowels that exist in one but not the other.

Table 1

Turkish vowel system, left, and Standard Southern British English, right.

We can safely assume that vowels shared between the two languages, especially cardinal vowels, will be produced with a similar F1 and F2 placing them in approximately the same region of the vowel space. These would be /i, ɑ, u/ primarily, though possibly with the addition of the mid-vowels /ɛ, ɔ/. These vowels will very likely differ slightly in quality, and could pose a problem to L1 English speakers as they may automatically categorize what they hear as a token of that vowel and not alter their productions. The novel Turkish vowels may be treated in one of two ways. First, as novel segments they will be easier to acquire due to their perceptual salience and we can expect to see greater imitation of these vowels. Second, they could be collapsed into the nearest English phoneme and this will interfere with acquisition of the contrasts. Which phoneme they are considered a token of is a little difficult to pin down. According to de Jonge et al. (2022) American English speakers have been shown to categorize [œ] as somewhere between [ɛ] and [ʊ]. We can posit that /y/ will be perceptually assimilated to SSBE [u], and /ɯ/ could be spread across /ʊ/, /u/, or /ʌ/ (de Jonge et al., 2022).

Methodology

Materials

Two lists of words, one English and one Turkish, constructed to elicit the vowels of each language are used. The English word list contains ten words designed to elicit the ten

monophthongs in SSBE with one word per vowel in CVC structure (Table 2). Included are six distractor words for a total of sixteen words which are displayed on screen in random order for the participants to read.

The Turkish word list contains eight words in CVC structure with six of the words forming minimal pairs in rounding (Table 3). We have used monosyllabic words to avoid Turkish vowel harmony and also to be in keeping with the English word list. Eight distractor words are included to bring the list to sixteen words in keeping with the English list.

Participants

All participants and model talkers in the study are required to be over the age of eighteen years old, with no diagnosed hearing or speech difficulties.

Monolingual English Speakers: A total of 40-60 monolingual Standard Southern British English (SSBE) speakers will partake in the study. The only additional requirement for participation is zero prior experience with Turkish, which here means that they have not had any formal or informal instruction. Exposure to, and experience with, other languages will be collected on the background questionnaire in the event that it could affect results.

Model Talkers: A single speaker of American/Canadian English provides the stimuli for the first phase of the experiment testing imitation levels in English.

Four native Turkish speakers, two female and two male, provide the stimuli for the second phase. To control any influence stemming from dialectal variations we will attempt to recruit speakers from a single dialectal area. Multiple model talkers are used in this portion of the experiment to encourage imitation of the general features of Turkish – not of one speaker specifically. The word list will be read in full by all speakers and divided equally between the model talkers in the shadowing portion, to ensure participants are exposed to all model talkers equally.

Procedure

The main experiment is divided into two phases, with the first phase assessing the imitation level of our participants in their native language, English. The second phase then proceeds to test their imitation in Turkish.

Phase 1: Participants will complete two blocks (equaling 3 repetitions) of a shadowing task, where they will hear the English model talker, one word at a time, and then repeat the word aloud. Results from this phase will then be used to divide the participants into one of two analytical groups: high imitators and low imitators. To be considered a "high imitator" a participant must show a greater than 60% success rate – meaning that in more than 60% of tokens they display clear imitation – with the remainder of participants going into the other group.

Phase 2: Participants will complete two blocks (equaling 4 repetitions) shadowing the Turkish word list. Like Phase 1, there will be one word per Turkish vowel (8 words total). In total there will be 64 tokens of Turkish shadowing, or 8 tokens per vowel for each participant. Unlike Phase 1, participants will not be reading the Turkish word list aloud first and will only complete a shadowing task.

Expected Findings

Due to the results of previous research (Lewandowski & Jilka, 2019; Wade et al., 2021) we can expect that level of imitation will carry over from L1 to L2. Likewise, just as in research by Spinu et al. (2018) we can expect high imitators to have a perceivable advantage over the lower imitators in their productions of Turkish vowels.

This will not apply to all vowels equally, as found by Kartushina et al. (2016) we expect that proximal vowels will show less imitation and may be entirely supplanted by an L1 vowel. In particular we expect this for Turkish /i, e, u/ which exist in SSBE and are likely to be categorized as tokens of these vowels.

References

Coles-Harris, E. H. (2017). Perspectives on the motivations for phonetic convergence [e12268 LNCO-0713.R1]. *Language and Linguistics Compass*, *11*(12), e12268. https://doi.org/https://doi.org/10.1111/lnc3.12268

de Jonge, K., Maxwell, O., & Zhao, H. (2022). Learning on the field: L2 turkish vowel production by l1 american english-speaking ngos in turkey. *Languages*, *7*. https://doi.org/10.3390/languages7040252

Jiang, F., & Kennison, S. (2022). The impact of l2 english learners' belief about an interlocutor's english proficiency on l2 phonetic accommodation. *Journal of Psycholinguistic Research*, *51*(1), 217–234. https://doi.org/10.1007/s10936-021-09835-7

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2016). Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics*, *57*, 21–39. https://doi.org/https://doi.org/10.1016/j.wocn.2016.05.001

Lewandowski, N., & Jilka, M. (2019). Phonetic convergence, language talent, personality and attention. *Frontiers in Communication*, *4*. https://doi.org/10.3389/fcomm.2019.00018

Sabev, M. (2019). *Spectral and durational unstressed vowel reduction: An acoustic study of monolingual and bilingual speakers of bulgarian and turkish* [Doctoral dissertation, University of Oxford]. University of Oxford.

Spinu, L. E., Hwang, J., & Lohmann, R. (2018). Is there a bilingual advantage in phonetic and phonological acquisition? the initial learning of word-final coronal stop realization in a novel accent of english. *International Journal of Bilingualism*, *22*(3), 350–370. https://doi.org/10.1177/1367006916681080

Wade, L., Lai, W., & Tamminga, M. (2021). The reliability of individual differences in vot imitation [PMID: 32772791]. *Language and Speech*, *64*(3), 576–593. https://doi.org/10.1177/0023830920947769

Zając, M., & Rojczyk, A. (2014). Imitation of english vowel duration upon exposure to native and non-native speech. *Poznan Studies in Contemporary Linguistics*, *50*(4), 495–514. https://doi.org/doi:10.1515/psicl-2014-0025

Table 2

English word list with corresponding target vowels

Word	Vowel (IPA)
heed	/i/
hid	/ɪ/
head	/ɛ/
had	/æ/
hod	/ɔ/
hawed	/o/
hood	/ʊ/
who'd	/u/
hud	/ʌ/
heard	/ɚ/

Table 3

Turkish word list with corresponding target vowels

Word (Orthography)	Word (IPA)	Vowel (IPA)
diş	/diʃ/	/i/
düş	/dyʃ/	/y/
deş	/deʃ/	/e/
döş	/dœʃ/	/œ/
kaş	/kaʃ/	/a/
dış	/dɯʃ/	/ɯ/
duş	/duʃ/	/u/
koş	/koʃ/	/o/

Prosodic Multimodal Perception in Second Language Sentences Disambiguation

Haohan Zhang[1], Jovana Radisavljevic[1] & Solveigh Janzen[2]
[1]Université Paris Cité, [2]Universität zu Köln

The human language comprehension system encompasses both linguistic and non-linguistic information (Tanenhaus et al., 1995). Numerous studies have illustrated that prosody aids listeners in disambiguating sentence structures (Snedeker & Trueswell, 2003). Additionally, the integration of visual and auditory cues has proven advantageous for language comprehension (Nakamura et al., 2012). This integration sheds light on an intriguing aspect of multimodal function in human language perception. For second language (L2) speakers, research indicates that prosodic cues in speech play a role similar to that in native speakers, assisting L2 learners in sentence interpretation (Dekydtspotter et al., 2008). However, their multiple sources performance may be limited, particularly in real-time comprehension (Nakamura et al., 2020).

The current research proposes a multimodal peception study inspired by Nakamura et al.'s (2020) research. It integrates prosodic cues and visual contextual stimuli to explore two main aspects: 1) Whether L2 English learners with diverse L1 backgrounds (French or German) exhibit different performances in perceiving syntactically ambiguous English sentences; and 2) Whether supplementary visual cues influence comprehension performance.

The experiment comprises two parts: 1) An auditory-only perception experiment and 2) An auditory-visual multimodal eye-tracking perception experiment. Following a Latin-square design, each part consists of items containing 24 syntactically ambiguous sentences. These sentences vary in two distinct intonation patterns: H* L- and the combination of H* (!H*) L- (as noted in E_ToBI). The sentences selected for testing involve syntactic ambiguity, such as (1). In this case, the noun phrase (NP) is ambiguous between serving as a modifier for the preceding adjective "history" (e.g., 1a) or as a modifier for the noun following it (e.g., 1b). The ambiguous modifier is marked by different accents, either H* or H* L-. All auditory stimuli are recorded by a male native speaker of American English.

(1a) Paul is an [American HISTORY $_{\text{H* L-}}]_{\text{NP}}$ teacher]]. (1b) Paul is an [American [HISTORY$_{\text{!H*}}$ teacher]$_{\text{NP}}$].

Twenty L2 English learners will be recruited for each L1 background, all of whom are monolingual native speakers of French or German, possessing an intermediate-advanced English proficiency level (CEFR B2-C1). In the experiment, participants will listen to sentences played audibly. In the first part, they will view an orthographically presented sentence with indices of the syntactic structure for two different interpretations (e.g., 1). Using a mouse, participants will click on the corresponding sentence they heard. In the second part, there will be only two visual interpretations of the ambiguous sentence on the screen, and participants will be asked to focus on the corresponding interpretation, fixing their eyes on the chosen visual stimulus. In each trial, participants' eye movements will be recorded using EyeLink 1000 (SR Research) at a sampling rate of 1000 Hz. Participants underwent a calibration procedure before the experiment started. The analysis will include the correct proportion of looks, the reaction time until the first correct look.

We hypothesize that at the auditory perceptual level, when the prosodic cues aiding sentence disambiguation align more closely with the prosodic features of participants' L1, they will perform better than the other group. For instance, a French L1 learner of English might be more sensitive to the boundary tone when it plays a role in phrases' segmentation. However, both groups may encounter challenges in the eye-tracking disambiguation task due to its reliance on visual perception and its connection to grammar.

References:

Dekydtspotter, L., Donaldson, B., Edmonds, A., Fultz, A., and Petrush, R. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English-French learners. *Stud. Sec. Lang. Acquis*. 30, 453–480. http://doi.org/10.1017/s0272263108080728

Nakamura, C., Arai, M., and Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition* 125, 317–323. https://doi.org/10.1016/j.cognition.2012.07.016

Nakamura, C., Arai, M., Hirose, Y., & Flynn, S. (2020). An extra cue is beneficial for native speakers but can be disruptive for second language learners: Integration of prosody and visual context in syntactic ambiguity resolution. *Frontiers in Psychology*, *10*, 2835. https://doi.org/10.3389/fpsyg.2019.02835

Snedeker, J., and Trueswell, J. (2003). Using prosody to avoid ambiguity: effects of speaker awareness and referential context. *J. Mem. Lang*. 48, 103–130. https://doi.org/10.1016/s0749-596x(02)00519-3

Tagliani, M., & Redolfi, M. (2023). Tracking Eye Movements as a Window on Language Processing: The Visual World Paradigm. *Encyclopedia*, *3*(1), 245-266. https://doi.org/10.3390/encyclopedia3010016

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., and Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–1634. http://doi.org/10.1126/science.7777863

Van Ommen, S., Boll-Avetisyan, N., Larraza, S., Wellmann, C., Bijeljac-Babic, R., Höhle, B., & Nazzi, T. (2020). Language-specific prosodic acquisition: a comparison of phrase boundary perception by French- and German-learning infants. *Journal of Memory and Language*, 104108. https://doi.org/10.1016/j.jml.2020.104108

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, *99*(2), B63-B72.

The effect of training on vowel perception and production in non-native speakers: The case of learners of French

Thomas BESSE, Tacita BLACK, Alexander ILGERT, Viktoriia VERSHNIAK

As previous studies have shown (Escudero *et al*., 2009), there are several possible cognitive statuses of sound categorisation and perception processes in L2 learners. It is indeed uneasy to distinguish the sounds that are non-contrastive in one's phonological spectrum. In the current study, we aim to focus on French anterior mid-vowels such as /e/ in *dé, clé, allez,* /ɛ/ in *crème, beige, tête,* /œ/ in *œuf, bœuf, cœur and* /ø/ in *veut, adieu, sérieux.* In contrast to English (Received Pronunciation), where only two of these vowels (/e/ and /ɛ/) coexist as allophones, French presents a phonemic contrast.

This study plans to employ phonetic training to investigate the extent to which perceptual mappings of these contrasts are malleable, and whether an effect of training on perception may extend to production. Indeed, high variability phonetic training (HVPT), in particular, has previously been found to increase learners' perception of non-native sounds (Barriuso & Hayes-Harb, 2018), assisting the building of new phonological categories (Melnik & Peperkamp, 2020). This will be manipulated as a comparative factor in this study, along with musical training, which may emerge as a noteworthy factor contributing to advanced perception and, consequently, production abilities. Approaches such as the Suzuki method or the Orff method, which include musical instrument training and sensibilisation in order to recognise even small differences in pitch, tempo and dynamics have shown promising results (Barbaroux *et al*., 2019; Flagge *et al*., 2021; Incognito *et al*., 2022). The commonality of the processing of musical impressions and prosodic elements of speech in the right brain hemisphere can be seen as an indication of this.

The proposed study will address the following research questions:

1. How do anglophone learners of French categorise anterior mid-vowels present in French pre- and post-training?
2. How do musical and phonetic training – respectively and comparatively – effect the perception and production of these vowels?
3. What is the relationship between perception and the production of these vowels?

The study will recruit L1 English learners of French as well as L1 French speakers to participate in the experimental procedures. The experimental procedures will take place in three stages and will evaluate both perception and production:

I. Pre-test stage:
   A ***phonetic categorisation task*** (perception): 20 stimuli spanning over a continuum with two degrees of freedom (aperture, labialisation) will be used to determine categorical boundaries.
   A ***reading task*** (production): Formants will be analysed and retrieved in a graph marking position in anteriority, aperture, and coloured depending on the degree of labialisation. For each task, an identification function will be calculated based on the performance of L1 French speakers.
II. Training stage:
   Participants will complete multiple phonetic (of which the conditions of speaker and context variability will be manipulated) and musical training sessions. Groups are envisioned as followed: a) No training received, b) High Variability Phonetic Training (HVPT), c) Musical training, d) Both.
III. Post-test stage: The same tests (I) will be administered to investigate the use of the target structure post-training. The results will be compared with those of Step I.

We predict that both the phonetic and the musical training will show effects on the perception and production tasks and improve the results of participants who received at least one of the mentioned treatments.

References

Aliaga-García, C., & M., Joan C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. *Recent Research in Second Language Phonetics/Phonology: Perception and Production*.

Barbaroux, M., Dittinger, E., & Besson, M. (2019). Music training with Démos program positively influences cognitive functions in children from low socio-economic backgrounds. *PloS One*, 14(5). https://doi.org/10.1371/journal.pone.0216777

Barriuso, T. A., & Hayes-Harb, R. (2018). High Variability Phonetic Training as a Bridge from Research to Practice. *CATESOL Journal*, 30(1), 177-194.

Escudero, P., Boersma, P., & Hamann, S. (2009). The linguistic perception of SIMILAR L2 sounds. In *De Gruyter eBooks* (pp. 151–190). https://doi.org/10.1515/9783110219234.151

Flagge, A. G., Neeley, M. E., Davis, T. M., & Henbest, V. S. (2021). A preliminary exploration of pitch discrimination, temporal sequencing, and prosodic awareness skills of children who participate in different school-based music curricula. *Brain Sciences*, 11(8). https://doi.org/10.3390/brainsci11081053

Incognito O., Scaccioni L., & Pinto, G. (2022). The impact of a music education program on meta-musical awareness, logical-mathematical, and notational skills in preschoolers. *International Journal of Music Education*, 40(1), 90-104.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology,* 54(5), 358–368. https://doi.org/10.1037/h0044417

Melnik, G. A., & Peperkamp, S. (2021). High-Variability Phonetic Training enhances second language lexical processing: evidence from online training of French learners of English. *Bilingualism: Language and Cognition,* 24(3), 497-506. https://doi.org/10.1017/S1366728920000291

Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012, May 23). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience*, 219, 111-9. https://doi.org/10.1016/j.neuroscience.2012.05.042.

Patricia CAMUS (Université Paris Cité)
Jessy Ans GONZAGA (Université Paris Cité)

Exploring the Influence of Personality Traits on VOT Phonetic Imitation: The Case of L2 Philippine English Speakers and American English Model Stimuli

1. Research Questions

For this paper, we are interested in answering the following research questions: *How does Voice Onset Time (VOT) shift across participants (L2 speakers of Philippine English) after exposure to model speech (American English), and which personality traits are correlated with a speaker's likelihood to imitate?*

2. Background

English is one of the official languages in the Philippines, alongside Filipino which is the standardized register of the language, Tagalog. English and Tagalog have been in contact and used together in the country since English was introduced in schools by the American colonizers in 1898. Majority of Filipinos have Tagalog or another Philippine language as an L1 and learn Philippine English as their L2 in schools as it is the medium of instruction in most school subjects. Due to this influence of L1, regarding the Voice Onset Time (VOT) of consonants stops, there is a clear influence of Tagalog in Philippine English (Lesho, 2017). Philippine English voiced stops such as /b/, /d/, and /g/ are characterized by having long negative (pre-voicing) VOT, a characteristic that is also present in Tagalog. Conversely, American English has a zero or short-lag when it comes to voiced stops (Docherty, 1992). For voiceless stops, Philippine English exhibits positive VOTs. This is similar to American English voiced stops that also have a positive VOT. However, while Philippine English voiced stops are unaspirated, American English voiced stops are aspirated which results in very long positive VOTs. Given these characteristics of Philippine English, we have chosen to focus on VOT in the productions of L2 speakers of Philippine English.

Among the many personality models available, we will use Costa & McCrae's (1992) Five Factor Model of Personality. It includes five factors which are each broken up into two values: Openness to Experience (inventive/curious vs. consistent/cautious), Conscientiousness (efficient/organized vs. extravagant/careless), Extroversion (outgoing/energetic vs. solitary/reserved), Agreeableness (friendly/compassionate vs. critical/rational), and Neuroticism (sensitive/nervous vs. resilient/confident).

3. Hypothesis

We are particularly interested in looking at personality traits as we intend to contribute to the inquiry about *who* the high imitators are, and what individual characteristics they possibly have that lead to higher likelihood of imitation, accommodation, or convergence. Taking into account previous studies about the influence of personality on likelihood of accommodation, we propose the following hypothesis:

- L2 speakers of Philippine English with the following personality traits are hypothesized to exhibit similar VOTs to their conversation partners after exposure: high Agreeableness, high Openness, high Neuroticism, low Extraversion, and low Conscientiousness. This implies that these personality traits lead to higher likelihood of accommodation.

4. Methodology

In order to investigate our research questions, we propose to conduct two experiments, namely an imitation task and a quasi-spontaneous task-oriented interaction. These two tasks are done on different days. The participants will be asked to take two tests prior to the tasks: (1) a phonetic talent test by Dogil & Reiterer (2009) and (2) a personality assessment test by Maples-Keller et al. (2019). We plan to make phonetic talent, the ability to grasp fine phonetic detail within a conversation and re-use it shortly after to adapt to their conversation partner, a controlled variable. Therefore, we plan to choose participants that have similar phonetic talent scores. All participants must be L1 speakers of Tagalog and L2 speakers of Philippine English who were born and raised in the Philippines. For the imitation task, we plan to only get female participants to make gender a controlled variable. According to Babel (2014), female speakers tend to imitate more than male speakers. In addition to this, studies have revealed that there exists a gender difference in VOT production wherein female speakers produce longer VOTs than male speakers (Swarts, 1992 & Robb, 2005, as cited in Kim, 2019). Although this gender effect in VOT has not been investigated in Philippine English, it is better to make this a control variable, as our main concern for this research is the effect of personality traits. In order to account for the gender effect, we plan to only recruit female speakers in the first task. On the other hand, to not dismiss the possibility of finding different results, we will recruit male participants for the quasi-spontaneous task-oriented interactions. However, they will be in controlled groups. This will be done in order to avoid possible sociolinguistic effects in accommodation regarding speaking with the opposite sex. For the imitation task, we plan to recruit 30 female speakers aged 18 to 35. On the other hand, for the quasi-spontaneous task-oriented interaction experiment, we plan to make 2 gender control groups: group 1 (15 female speakers speaking with each other) and group 2 (15 male speakers speaking with each other). For the American English stimuli, we plan to hire one American male adult who speaks American English as his mother tongue to participate in producing stimuli in the imitation task and interacting with participants in the quasi-spontaneous task.

4.1 Imitation Task

We plan to do the imitation task experiment in 3 steps. The first step is a b*aseline production task* where participants will be tasked with generating a list of 60 initial target words for both voiceless stops /p/, /t/, /k/ and voiced stops /b/, /d/, /g/ within the specified carrier sentence “say __ again”. The total number of words is 120 and will be chosen based on lexical frequency per quartile from the CELEX lexical database of English. The total number of words was calculated in the following equation: 4 quartiles of lexical frequency x 3 places of articulation of stops (bilabial, cooronal, and velar) x 5 examples of words per quartile = 60 items for voiceless and 60 items for voiced. The

second step is the *perceptual learning task* or exposure block. Here, the participants will be exposed to a recorded first-person narrative delivered by the male American English speaker hired for our experiment, incorporating the same set of words in a randomized order. The third step of the task is a *post-exposure production task* where participants will be asked to reproduce the identical word list, encompassing both /p/, /t/, /k/ and /b/, /d/, /g/, but in a different randomized order.

4.2 Quasi-spontaneous task-oriented interactions

For the quasi-spontaneous task-oriented interactions, we have chosen a map task and a picture-matching game ("spot the difference"). The establishments in the map task as well as the different items in the picture-matching game will consist of the target words we are interested in. We will include 30 target words for both voiceless stops /p/, /t/, /k/ and voiced stops /b/, /d/, /g/. For example, we can include words like *park, tea shop, dentist, ball, dress, girl,* etc. as target words, whose VOTs we will measure. Participants will have to accomplish two map tasks and two picture-matching games within five minutes each (maximum of 20 minutes per participant). Each subject will interact with an American English speaker as we try to see whether VOT imitation can happen in real-time.

5. Analysis

A two-step modeling approach will be done for our study, with slight differences on the two experiments. First, we need to calculate the disparity between VOT productions of participants. For the imitation task, we will calculate the disparity between pre-narrative and post-narrative normalized VOT values for each target word using PRAAT. On the other hand, for the quasi-spontaneous task-oriented interactions, there are no pre- or post-exposure blocks. As such, we plan to explore conversation chunks at various granularities, such as halves, thirds, quarters, etc., and finer granularity at the turn level. The second step of our approach is to analyze the influence of personality traits on the degree of VOT shift. Finally, we will compare the results of the two experiments.

6. References:

Babel, M. (2014). Novelty and social preference in phonetic accommodation. University of British Columbia. Laboratory Phonology 2014; 5(1): 123 - 150.

Baker, R., & Hazan, V. (2011). DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. Behavior research methods, 761–770.

Costa, P. T., Jr., & McCrae, R. R. (1992). NEO PI-R professional manual. Odessa, FL: Psychological Assessment Resources, Inc.

Docherty, G. (1992). The timing of voicing in British English obstruents. Berlin.

Dogil, G., and Reiterer, S. M. (eds.). (2009). Language Talent and Brain Activity, Vol. 1. Trends in Applied Linguistics. Berlin: De Gruyter. doi: 10.1515/9783110215496

Kim, S.K. (2019). Gender differences in voice onset time perception. The Journal of the Acoustical Society of America. DOI: https://doi.org/10.1121/1.5137586

Lesho, M. (2017). Philippine English (Metro Manila acrolect). Journal of the International Phonetic Association, 48(3), 357–370. https://doi.org/10.1017/S0025100317000548

Maples-Keller, J. L, Williamson, R. L., Sleep, C. E., Carter, N. T., Campbell, W. K., & Miller, J. D. (2019). Using item response theory to develop a 60-item representation of the NEO PI–R using the International Personality Item Pool: Development of the IPIP–NEO–60. Journal of Personality Assessment, 101, 4-5. DOI: 10.1080/00223891.2017.1381968

RESEARCH PROPOSAL

Cross-Linguistic Perceptual Variations in Rhotic Vowels between Native Mandarin and American English Speakers

Team members: Marie Huygevelde 22200088 (Université Paris Cité), Yunzhuo Xiang 22201420 (Université Sorbonne Nouvelle), Azim Uddin Ahmed (7404855, University of Cologne), Dong Yi 22212691 (Université Paris Cité)

Introduction This paper will be an endeavor to contribute to the understanding of cross-linguistic perceptual variations in rhotic vowels, bridging gaps between linguistic theories and empirical investigations across Mandarin Chinese and American English. Their two phonological systems present rhoticity implying a low F3 covering their entire length (Huang, 2010; Lee & Zee, 2014), and the values of F3 were shown to be similar for both in postvocalic position (Chen & Mok, 2021). Contributing to the debated questions on acoustic cues responsible for rhoticity perception (Heselwood & Plug, 2011), we will address the following questions via three experiments: To which extent the F3 lowering will be classified as rhotic? Does this perception differ depending on native language? Does attenuating F3 affect perception of rhoticity?

Methodology **Experiment 1:** A production study of English and Mandarin rhotic vowel will be conducted with American English (group 1) and Mandarin Chinese (group 2) monolingual adults (minimally 10 participants for each group) to give us a baseline for the perception experiment stimuli. The aim is to confirm the presence of a lowering in F3 in the acoustic signal and target potential differences to control for in the following experiments (especially, duration of both vocalic portion and rhoticity). Participants will utter minimal pairs of words with low back vowels [a] and [ɑ] (e.g. English: lop /lɑp/ - larp /lɑrp/; Chinese: 辣 /la/ - 烂儿 /lar/, etc) read in sentences realizing rhoticity in similar closed syllabic contexts, [CV(r)#p] context (“#” as syllable boundary), *She said again: “()”* (group 1*) and* 这个读“()”吧 *[ʈʂɤ kɤ tu “( )” pa]* (meaning: “This word pronounced as: “) (group 2) *producing* “烂儿吧” /lar#pa/ - larp /lɑrp/. Similarities are expected on acoustic cues in rhotic items: a lowering in F3.

Experiment 2: An identification task will investigate perception of acoustic rhoticity under two dimensions with two 11-step continua: (1) modulating the slope of F3 lowering (from 2500Hz to 1500Hz with a step each 100Hz), and (2) their attenuating F3 counterparts. Gathered from Exp 1, pairs of stimuli produced in similar consonantal contexts (Ex. larp - 烂儿) in both languages will be used as endpoints of continua. At each trial, participants will listen to stimuli of both languages and be asked to identify them as similar to one of the endpoints representing rhotic vs non-rhotic categories. Thus, all participants will encounter stimuli from both languages since rhoticity values are expected to be cross-linguistically similar in post-vocalic position (Chen & Mok, 2021), and in the two conditions involving a F3 lowering associated with an attenuation of the F3 intensity.

Experiment 3: In order to assess further the degree of both groups’ auditory sensitivity to the presence/absence of rhotic information in the acoustic signal, an eye-tracking method will be used. Extracted from perception stimuli, three sets will be created, following partly the methods in Beddor et al. (2013), including two with manipulation of the onset time of rhoticity in the vowel: (1) early onset rhotics [**CV**$_{early}$**r**]; (2) late onset [**CV**$_{late}$**r**]; (3) non rhotics [**CVC**]. To constrain the rhoticity spreading at the vocalic portion only, we will splice the stimuli to add the same initial consonant (taken from the non-rhotic sequence). Here, temporal variation of the onset is also used to avoid stimuli habituation. In front of a screen, participants with Eyelink headgear will face a fixation cross then a pair of written names of the audio stimuli (vocalic early rhotics vs non-rhotics; vocalic late rhotics vs non-rhotics). Before listening to audio stimuli embedded in the same sentences as Exp 1, they will be asked to look at the written word corresponding to the heard one. Due to difficulties to represent graphically the words, they will be presented in rectangular frames providing boundaries to calculate fixations latencies between the fixation cross and the target word.

Predictions: The productions in **Exp 1** are expected to be similar in terms of formant patterns indicating rhotic quality but varying in duration of the vocalic and rhotic portions between and within groups. The latter will be analyzed by a t-test and averaged to the mean in case of statistical significant differences. For the **Exp 2**, we expect to see similar identification patterns between Mandarin and English speakers, with attenuated F3 stimuli more perceived as rhotic because of the F-rho dominating effect (Heselwood & Plug, 2011). Dividing the response rate along a 33-66% threshold, we foresee an earlier shift in identification for the attenuated continuum. ANOVA will be used to analyze crossed effects of F3 attenuation, native language and F3 lowering. In **Exp 3**, visual fixation latencies will be measured from the onset of the stimuli until the point the eye gaze reaches the frame boundaries of the visual correct item. Analyzed via a linear mixed-effect model and taking as baseline the latencies in non-attenuated [CVC - CV$_{earlyr}$] trials, we predict aligning results with Exp 2 showing shorter latencies with F3 attenuated target stimuli. Focusing particularly on non-rhotic vs rhotic fixations responses, we expect overall shorter latencies for both [CVC - CV$_{earlyr}$] and [CVC - CV$_{later}$] attenuated trials compared to non-attenuated ones. We do not hypothesize sensitivity inducing significant differences in [CV$_{earlyr}$ - CV$_{later}$] trials latencies within both conditions. Moreover, reflecting perception patterns, we predict an effect of position in the continua with differentiating stimuli at the extremes faster than the ones closer to the threshold boundaries, i.e. stimuli pairs at steps 1 vs 9 triggering shorter latencies than at steps 3 vs 7.