



Open Library of Humanities

Prosodic location modulates listeners' perception of novel German sounds

John H. G. Scott*, School of Languages, Linguistics, Literatures and Cultures, University of Calgary, Calgary, AB, Canada, john.scott@ucalgary.ca

Isabelle Darcy, Department of Second Language Studies, Indiana University, Bloomington, IN, USA, idarcy@indiana.edu

*Corresponding author.

Interaction of sounds on the melodic tier (segments) with prosodic and phonotactic structure (syllabic context) in cross-language perception is not explicitly addressed by models of second language phonology (e.g., Perceptual Assimilation Model: Best, 1995). At initial stages of foreign language exposure, learners rely on position-specific phonetic detail more than native speakers or advanced learners, thus mappings according to prosodic and phonotactic context are a crucial factor in early interlanguage phonological development. In a perceptual assimilation experiment, we manipulated syllable position (onset vs. coda) and phonotactic complexity (simple vs. complex codas) in phonotactically similar languages for auditory presentation of six German obstruents (i.e., familiar [h k ʃ] and novel [ç x p̄f]) to native speakers of American English who had no previous exposure to German. By means of weighted proportions (Park & de Jong, 2008) and overlap scores (Levy, 2009), we found that [h k ʃ p̄f] mapped categorically to English orthographic categories <h>, <k>, <sh>, and <f>, respectively, in all positions, whereas the novel fricatives [ç x] exhibited distinct mapping patterns from each other, from other sounds, and according to syllable position. These results demonstrate profound influences of both low-level prosodic and phonotactic contexts on perceptual assimilation of novel sounds.



1. Introduction

One of the most important advances in cross-language speech perception has been to show that sounds from a different phonological system are interpreted by listeners in terms of their native language (L1) categories (e.g., Best, 1995; Polivanov, 1931; Strange, 2011). A vast number of studies have endeavored to establish a typology of categorization patterns by which nonnative sounds are associated to native categories. For example, a prominent model by Best (1995), the Perceptual Assimilation Model (PAM), suggests that through the mechanism of perceptual assimilation, nonnative sound exemplars are perceived as good or less good exemplars of native categories—depending on their gestural similarity with that native category. Other models (e.g., Speech Learning Model: Flege, 1995; Flege & Bohn, 2021; Flege & MacKay, 2004) have proposed similar mechanisms such as perceptual equivalence classification, thereby seeking to explain some of the difficulties that second language (L2) learners experience when distinguishing some contrasts in their L2 or why L2 learners produce some L2 sounds using their L1 category definitions.

Such perceptual mapping data are impactful because they can be used to predict perceptual difficulties between nonnative sounds as a function of how they relate to native categories. For example, numerous studies have documented the well-known difficulty exhibited by L1 Japanese speakers perceiving the difference between American English /ɹ/ and /l/ (e.g., *rock* vs. *lock*). This difficulty in discrimination is commonly understood to arise from the mismatch between the two American English categories /ɹ/ and /l/ and the single Japanese central approximant /r/ to which they both are associated in cross-language perception (e.g., Brown, 1998, 2000). This and other categorization types are summarized in §1.1.

Although predicting discrimination performance from categorization types is not always straightforward, recent approaches have shown that overlap scores can predict performance very well (Daidone, Kruger, & Lidster, 2015, 2019; Levy, 2009), especially in the case of consonants. However, a major limitation of most models of cross-language perception of segments (such as the prominent PAM) is that they do not yet explicitly include prosody as a modulating factor, and they do not often consider the detailed phonetic factors pertaining to the environment of the sound, such as coarticulation with adjacent segments or prosodic position. To take a step in this direction, the present study includes the position of consonantal segments within the structure of a syllable (a low-level prosodic domain) as a factor of interest in perceptual assimilation. While acknowledging the importance of higher prosodic domains such as the word and sentence levels, here we necessarily focus on subsyllabic structure in conjunction with phonotactic context.

In general, perception studies often consider adjacent segmental context. It has long been established in speech perception that the quality of adjacent segments, in the form of coarticulation cues such as matched or mismatched formant transitions, can significantly modulate category

perception of speech sounds in discrimination and phoneme monitoring tasks, with measurable effects on accuracy or reaction time latencies (e.g., Levy & Strange, 2008; Otake, Yoneyama, Cutler, & Van der Lugt, 1996; Streeter & Nigro, 1979; Whalen, 1984, 1991). This empirical and methodological awareness is also reflected in numerous cross-language perception studies (e.g., Kilpatrick, Bundgaard-Nielsen, & Baker, 2019; Kilpatrick, Kawahara, Bundgaard-Nielsen, Baker, & Fletcher, 2018; Strange, Akahane-Yamada, Kubo, Trent, & Nishi, 2001), which lend support to the important generalization that the acoustic properties of a phoneme are sensitive to the context in which a particular instance of that phoneme appears. In short, the contexts in which a phoneme appears influence the forms of its context-local varieties—that is, its allophones. Furthermore, the relationships between a phoneme and its context-dependent allophones are learned and language-specific (e.g., Automatic Selective Perception [ASP], Strange, 2011), and so cross-linguistically arbitrary. For example, English maintains a clear phonemic distinction between alveolar /s/ and post-alveolar /ʃ/ (e.g., [sɪp] *sip* vs. [ʃɪp] *ship*). In contrast, coarticulation of /s/ in Icelandic, which lacks this distinction, with a following /k/ might yield [ʃ] as a viable variation of the phoneme, such as in ['i:s.lɛn.ʃky] *íslensku* 'Icelandic language,' a distribution analogous to that of velar [k] vs. palatal [ç] in English before back versus front vowels, respectively (e.g., /ku:l/ [ku:l] *cool* vs. /ki:l/ [çi:l] *keel*).

Unlike segmental coarticulation effects, cross-language perception studies that manipulate suprasegmental context or syllable position are less common. Shea and Curtin (2010) investigated the acquisition of L2 Spanish lenited obstruent allophones, showing that stress environment contributes to distributional learning. A separate strand of research in cross-language perception investigates how stress environment (pre- vs. post-tonic) and syllable position (onset vs. coda) interact with perception of English sounds by L1 Korean speakers. Lee and Cho (2006) investigated L1 Korean speakers' identification of L2 English consonants with International Phonetic Alphabet (IPA) labels, showing that prosody and stress modulate accuracy of identification: Consonants in onset position were identified more accurately than in coda positions. Cho and Lee (2007) demonstrated that L1 Korean speakers also exhibit more diverse mapping patterns for L2 English sounds to Korean categories in syllable codas and following unstressed vowels. Perceptual assimilation experiments, which incorporate both L2-to-L1 category mappings and goodness-of-fit (GF) ratings for each mapping, confirm that L1 Korean speakers presented with L2 English sounds exhibit both *less felicitous* and *more diverse* perceptual mappings of English sounds in syllable codas than in syllable onsets (Cho & Chung, 2010; Park & de Jong, 2008, 2017). In the pairing of L1 Korean and L2 English, these patterns of mapping interact both with phonotactic restrictions on distribution (i.e., not all consonants are equally permitted in Korean syllable codas) and neutralization of certain consonants in Korean codas that are contrastive in other positions (Park & de Jong, 2017). Crucially, none of the studies mentioned investigates an L1-L2 pairing in which the phonotactics of the coda in syllable templates are similar and no L1 neutralization process

limits the perception of consonants in coda position; the pairing of L1 English with German as a foreign language (FL) investigated in the present study meets both of these conditions.

Including syllable position as a modulating factor is important in order to obtain realistic predictions about perceptual difficulties which would reflect the complexity of the sounds' phonetic reality and of their local prosodic characteristics. The phonetic reality of sounds is affected by multiple factors. Just like phonetic environment, syllable position or phonotactic structure can all affect production (e.g., Broselow & Finer, 1991; Davidson, 2011; Eckman & Iverson, 2013); these factors can also affect perception (e.g., prosody in pronunciation instruction: Derwing & Munro, 2005; Jackson & O'Brien, 2011; perception by stress or syllable position: Park & de Jong, 2017; Winters, 2001; influence of adjacent segments: Kilpatrick et al., 2018; Kilpatrick et al., 2019). Consequently, examining how syllable position influences speech perception matters, as it can change how people perceive sounds. Too narrow a focus may indeed lead to erroneous conclusions about perceptual confusion patterns. For example, Moulton (1962, p. 31) confirms anecdotal evidence, which tells us that learners of German often confuse /x/ and /ç/ with /k/ in words like *Nacht* [naxt] versus *nackt* ('night' vs. 'naked'), perceiving and producing both words as [nakt] ('naked'). A narrow view of perceptual assimilation patterns of these sounds in only one position—such as the word-initial as in most traditional studies—would instead suggest that nonnative listeners map both sounds onto some kind of /h/, a conclusion at odds with the experience of many learners and educators. On the basis of codas only, most people might conclude that /x/ and /ç/ are perceived as a stop—a strange conclusion, because this does not happen in word-initial position. The potential implications of accurate predictions from richer models of cross-language perception range from identifying which specific words learners might find challenging (and why) to implementing more effective pronunciation instruction in entire curricula, taking prosodic factors into account (Derwing & Munro, 2005).

The present study considers position within the syllable as a potential modulating factor of perceptual assimilation of consonants. More specifically, we examine how American English listeners without any knowledge of German perceive unfamiliar German obstruent sounds such as [x] and [ç] by varying the subsyllabic constituent in which the phones appear (simple onsets, simple codas, and complex codas). This manipulation interacts with phonotactic sequencing constraints, because not every possible category of English sounds, onto which the German phones are expected to map, is acceptable in all positions. Therefore, this study contributes to documenting the interactions of syllable position, phonotactic influences, and segmental perception in cross-language speech perception.

1.1. Perceptual assimilation model

To put this study in theoretical context, in this section we review the PAM (Best, 1995) and previous studies that have examined the modulating factors of perceptual mapping patterns,

such as positional asymmetries in perception between subsyllabic constituents. Because our study also refers to phonotactic restrictions, we review previous investigations of phonotactics in L2 acquisition. We also provide a brief overview of German obstruents and discuss the restricted distribution of [h] in English and German phonotactics (Sections 1.2 and 1.3).

The Perceptual Assimilation Model (PAM; Best, 1995) and its L2 extension (PAM-L2; Best & Tyler, 2007) outline the nature and mechanisms of speech perception and category assimilation of nonnative language sounds in order to predict the difficulty of learning them. The PAM targets foreign speech perception by naïve listeners, but it hints at the potential to expand its principles into L2 acquisition. The subsequent PAM-L2 makes this expansion explicit, maintaining that fundamental differences exist between naïve listeners in the FL context and L2 learners—namely, status as a learner, linguistic environment, and influence of developing L2 categories on the common interlanguage (IL) phonological space.

The PAM(-L2) claims that FL or L2 segments are perceived according to how similar they are to the L1 segment(s) closest in proximity in the common phonological space (i.e., goodness-of-fit: GF), and whether a FL or L2 sound falls within the bounds of the speaker's "native phonological space" (e.g., Do L1 English speakers recognize isiZulu clicks as speech sounds at all, and can they learn to?). The PAM(-L2) measures "similarity" in terms of properties of the articulatory gestures and their proximity in the universal phonetic domain defined by the structure and potential perturbations of the vocal tract. In short, the PAM(-L2) addresses patterns and mechanisms by which novel phones from an unfamiliar language or a L2 are mapped (or not) onto L1 phonemic categories and provides a taxonomy of L2-L1 contrast relationships.

According to the PAM(-L2) taxonomy, FL/L2 contrasting sounds may be assimilated to two separate L1 categories (Two-Category: TC Type), to a single L1 category with differential GF (Category Goodness: CG), or to a single L1 category with equivalent GF (Single Category: SC). Additionally, FL/L2 pairs may both fail to assimilate to any existing L1 category, but still fall within the native phonological space (Both Uncategorizable: UU), or one FL/L2 phone may assimilate to a L1 category while the other does not, despite being within the native phonological space (Uncategorized vs. Categorized: UC). If neither FL/L2 phone falls within the native phonological space, no assimilation to L1 categories occurs (Nonassimilable: NA). Each type is hypothesized with an associated level of difficulty for naïve listeners and L2 learners ranging from "poor" to "very good" (Best, 1995, pp. 194–197; Faris, Best, & Tyler, 2018).

An important consideration pertains to how well these categories (or assimilation types) fare in predicting actual discrimination difficulties. Levy (2009) and Daidone et al. (2015, 2019) have shown that overlap scores (a measure of perceptual similarity of nonnative sounds to each other) are a much more accurate predictor of such difficulties. Overlap scores were always better predictors than categorization types (Daidone et al., 2019, Slide 31). Even though categorization types were acceptable predictors with consonants in Daidone et al. (2019), overlap scores were

an even better predictor of discrimination than nonnative to L1 categorization. This is why, in this study, we do not generate categorization types, but instead a weighted proportion (Park & de Jong, 2008) and an overlap score (Levy, 2009).

A number of studies have examined how perceptual assimilation patterns are modified by different factors. Syllable position (Cheng & Zhang, 2015; Park & de Jong, 2008, 2017; Scott, 2019; Sheldon & Strange, 1982) and segmental context (e.g., Levy & Strange, 2008; Scott, 2019) have been shown to influence the perceived phonetic similarity, and consequently, the mapping. Strange, Akahane-Yamada, Kubo, Trent, Nishi, and Jenkins (1998) manipulated the sentence context and found overall less clear mapping patterns in sentences compared to an isolated word condition. Word stress has also been found to impact perceptual patterns (Park & de Jong, 2017; Rose, 2010; Rose & Darcy, 2011 [unpublished dataset]). In addition to syllable position, segmental context, and word stress, several more factors have been shown to affect the two main determinants of perceptual assimilation (i.e., phonological structure and phonetic detail) on discrimination, identification, and cross-language mapping tasks. These include lexical status (word-nonword, Yoshida & Hirasaka, 1983), subjective frequency (de Jonge, 1995; Flege, Takagi, & Mann, 1996), and metalinguistic factors such as orthographical availability of representation (Escudero, Hayes-Harb, & Mitterer, 2008; Flege et al., 1996; Smith & Kochetov, 2009).

In the present study, we interpret the perceptual assimilation results in terms of strength of mapping relationship (weighted proportions, Park & de Jong, 2008) and overlap scores (Levy, 2009) rather than according to the taxonomy of assimilation types presented by PAM (Best, 1995). Recent research by Daidone et al. (2015; 2019, Slide 31) demonstrates that classification by PAM assimilation type is highly sensitive to the different performance thresholds set between studies in the literature, which profoundly alters the predictions for discriminability made by that framework (see also Faris, Best, & Tyler, 2018, p. 17). Even though categorization types were good predictors for consonants in Daidone et al. (2019), overlaps scores (Levy, 2009), which are not subject to arbitrary performance thresholds, were even stronger predictors of discrimination results for novel FL/L2 sounds. Perceptual assimilation experiments have the design advantage that many tokens can be included, which makes it relatively easy to manipulate syllable position conditions. This advantage must be exploited in moderation to avoid the drawback of increasing the number of conditions, which also multiplies task length in a balanced design. In summary, we employ the perceptual assimilation task for its design advantages; in addition, we analyze the results in terms of overlap scores as well as weighted proportions of the mappings, because perceptual similarity of nonnative sounds to each other is a better predictor of discrimination than categorization of nonnative sounds to the L1 inventory.

Perceptual asymmetry of segments on the basis of syllable structure constituency is well documented in L2 phonology, particularly for learner populations whose L1 has more phonotactically restricted syllable codas than the L2 (e.g., L1 Mandarin L2 English, Cheng &

Zhang, 2015). Park and de Jong (2008, 2017; see also Cho & Chung, 2010; Cho & Lee, 2007; Lee & Cho, 2006) observed perceptual asymmetries between consonants in initial onset position, pre- and posttonic intervocalic positions, and final coda position in the performance of L1 Korean speakers with American English sounds. In their study, all coda consonants elicited poor GF and noisier mapping patterns across the board. They argue that both structural asymmetry by syllabic/prosodic environment and Korean-specific phonotactic restrictions and coda neutralization effects modulate listeners' perception. The present study seeks to disentangle these factors by investigating a phonotactically more similar pair of languages.

Asymmetry between onsets and codas is not limited to cross-language perception nor to languages with narrowly restricted syllable codas. Winters (2001) investigated perceptual salience cues for stops in English, confirming that listener sensitivity to the place of labial, coronal, and dorsal stops was significantly higher in syllable onsets than in syllable codas—that is, stops were more salient in onset position. This asymmetry between onsets and codas arises at the suprasegmental level as well. House (1996) presents evidence for asymmetric perception between tonal contours in syllable onsets at the beginning of the nucleus, where F0 change is accompanied by simultaneous rich, new spectral information, versus in the syllable rhyme (steady vowel state cues in the nucleus or the beginning of the coda), where F0 change is accompanied by relatively less new spectral information. According to his model of tone percepts, tone contours at the onset or beginning of the nucleus are perceived (in his terms, *coded*) as a level tone; tone contours in the nucleus or early in the coda are perceived as a contour tone, where tonal information is relatively more salient.

Levy (2009, p. 1139), citing other studies (e.g., Strange, Bohn, Trent, & Nishi, 2004), argues that differential sensitivity to contrasts by position is to be expected (see also Flege, 1995); however, subsyllabic constituent structure is not addressed in the PAM(-L2). The PAM is agnostic on the difference between onset versus coda positions and singleton consonants versus consonant clusters, while acknowledging that perception of L2 contrasts may be sensitive to the position in which the listening target occurs (Best, 1995; Best & Tyler, 2007). It is therefore important to document more widely how the interplay of phonotactics and syllable position impact perceptual mappings, so as to enable generalizations that could expand the models.

1.2. Phonotactics in L2 phonology

To discuss the role of phonotactics in L2 phonology, the definition of phonotactics itself warrants attention. Selkirk (1984) provides a formal definition of syllable phonotactics:

Given an autosegmental theory of the syllable, the phonotactic description of the syllable has at least three parts: (i) the characterization of possible syllable structures, (ii) the characterization of possible (or impossible) sequences on the melody tier, and (iii) the characterization of

possible associations between the two. Each of these is to be viewed as a set of well-formedness conditions. For a syllable to be ruled well formed, it must be well formed with respect to (i–iii). (p. 114)

Selkirk's first part of phonotactic description refers to syllable templates, such as CV (consonant-vowel) versus CVC, CCVC, CCVCC, and other syllable shapes. Her second part may broadly apply to any conditions governing permissible and impermissible sequences of segments in a given language. In Selkirk's third part lie references to sets of sounds that are (not) licensed to appear in a particular position in the structure of a syllable.

Investigations of phonotactics in L2 are broadly represented in the literature, and taken together, they demonstrate that the L1 phonotactic grammar can modify the perception of sounds in L2 depending on their position. Many studies have contrasted languages with different syllable or word structure—falling under Selkirk's (i) part (e.g., Cardoso, 2011; Cheng & Zhang, 2015; Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Kabak & Idsardi, 2007; Park & de Jong, 2008, 2017; Trapman & Kager, 2009). These studies demonstrate that underlying phonotactic differences between the FL/L2 and the L1 modulate perception of novel sounds. Notably, a few studies have examined the perceptual assimilation of illicit consonant clusters (e.g., Hallé, Segui, Frauenfelder, & Meunier, 1998; Kilpatrick et al., 2018; Segui, Frauenfelder, & Hallé, 2001)—falling under Selkirk's (ii) part. Studies of this type move beyond the single-segment analysis presented by the original PAM(-L2) to investigate the perceptual effects of segment adjacency. The present study differs from the first group by pairing two languages that have similar syllable structure templates: English and German. It resembles the second group by examining the behavior of clusters in codas. In addition, it differs from that group by considering a range of [t]-final coda cluster types, some of which are licit in both languages, some of which are illicit in both, and some of which are licit in the FL but include novel sounds. By focusing on languages that have similar syllable templates but different inventories of sounds, the present study investigates how familiar sounds are perceived with attention to acoustic detail in different positions in FL/L2 (e.g., Sheldon & Strange, 1982, for L1 Japanese patterns of [r]-[l] perception of acoustic differences by syllable position) and how novel sounds are perceived in relation to acoustically or articulatorily similar L1 sounds, either subject to phonotactic restrictions in both languages (i.e., [h] in codas) or licensed in both languages (i.e., [k] in codas before [t]). (See §1.3 for relevant details about the obstruents of interest.) By examining how these sounds interact with the position-specific constraint against [h] in codas, the present study falls—in part—under Selkirk's (iii) part.

Similar to the present study, Park and de Jong (2017) move beyond single-environment and segment adjacency contexts to investigate the effects of syllable position and stress environment on perceptual assimilation of consonants. In their study, L1 Korean speakers learning English

as a foreign language listened to nonwords constructed from the English consonantal segments /p b t d f v θ ð/ in the vocalic /a/ context. Their comparison of consonantal perception in VC, VC', and VC' environments shows extensive differences between mapping patterns in these positions versus those found in CV position by Park and de Jong (2008). However, the observed responses in coda positions indicate that L1 Korean Coda Neutralization plays a position-specific role in the perception of L2 English consonants. The specific contribution of syllable position versus phonotactic sequencing—that is, parts (i) and (ii) of Selkirk's definition—is difficult to pinpoint. In a recent psycholinguistic study, Kilpatrick et al. (2019) extend the concept of conceptual categories presented by PAM(-L2) in terms of segments to incorporate sequences of sounds. In their study, L1 Japanese listeners completed a forced choice perceptual assimilation task and an AXB discrimination task using VCV stimuli that either conformed with or violated phonotactic constraints of Japanese. Phonotactically illicit strings were assimilated to the nearest licit categories of Japanese. They found that the phonotactic properties of Japanese influenced L1 Japanese listeners in their perceptual assimilation of consonants that contrast in Japanese on the basis of adjacent vowels in VCV stimuli (cf. Selkirk's part [ii] of syllable phonotactics). Further, they found that L1 Japanese listeners exhibited higher accuracy and faster response times discriminating between conservatively licit sequences than between licit and innovative (atypical in Japanese) or illicit sequences. Their study calls for research to address the lack of specific testable predictions made by leading L2 phonology models with regard to acquisition of higher-order levels of phonological knowledge. These findings highlight the importance of phonotactics and position-sensitivity, both generally and language-specifically in perceptual assimilation and IL development. The experiment reported in our study remains constrained to (C)CVC(C) monosyllabic pseudowords, so as to focus investigation upon phonological constraints governing segmental sequences (ii) and the licensing of certain segments in specific syllable positions (iii).

1.3. German obstruents and restricted distribution of [h] in English and German

German and English are close historical relatives. Despite numerous historical sound shifts in both languages and famously prodigious lexical growth and replacement of historical words in English during recent centuries, both languages retain much of the phonemic and phonotactic character of West Germanic, their common ancestor. The shared historical origin of German and English makes them an ideal candidate pairing for study among natural languages due to their similarity in these areas.

Both German and English have a variety of obstruents ranging from bilabial [p b] to velar [k g] and placeless or glottal [h]. To avoid the issue of German final devoicing (i.e., neutralization of voice contrasts in syllable-final position), this study is limited to voiceless obstruents [h k ʃ], which are shared by German and English, and [ç x p̥f], which occur in German but are novel for

speakers of Standard American English without foreign language or dialect experience. Published research on fricatives in modern German is sparse and tends to focus on phonological mergers in specific local or regional dialects. Jannedy, Weirich, and Hemeke (2015) and Jannedy and Weirich (2017) quantify a perceptual merger of [ʃ] and [ç] in *Kiezdeutsch* ('Hood German,' see Wiese, 2006), a multi-ethnic urban dialect in the Kreuzberg area of Berlin, by means of a variety of spectral moment measurements in the noise of their frication (e.g., Center of Gravity, standard deviation, kurtosis, skewness) and discriminability by speakers of *Kiezdeutsch* versus speakers of another Berlin dialect and the dialect of Kiel. Kleber, Lowery, and Stegmaier (2018) investigate how speakers of Standard German distinguish the three-way /s ʃ ç/ contrast, employing similar acoustic measurements and identification judgements by native speakers on a three-way forced choice identification task with 13-step continua. Tronnier and Dantsuji (1993) investigate Japanese and German /h/ phonemes preceding /i/ in contrast with the German palatal /ç/ before /i/ or other vowels, reporting formant structure descriptions of each. Hall (2013) takes a phonological approach to describe how feature markedness motivated an unconditional historical merger of Middle High German palatal /ç/ and postalveolar /ʃ/ into alveopalatal /ç/ in certain modern Central German dialects.¹ To our knowledge, no published study has simultaneously investigated the acoustic properties of Standard German fricatives [ʃ ç x h], those employed in the present study, in a study comparable to Stevens's (1960) language-agnostic survey of fricative spectra by place of articulation. However, the studies cited here together establish the discriminability of three of them (i.e., [ʃ ç h]) in Standard German, although /ʃ/-/ç/ mergers obtain in some dialects.

For most American English speakers, [ç] occurs only due to coarticulation in certain clusters (e.g., *hue* [çju:]) and [x] may be encountered in other dialects (e.g., Scottish *loch* [lɒx] 'lake, sea inlet'; Yiddish English *chutzpah* ['xɒts.pə] 'arrogant audacity, impudence'); neither is phonemic. The American English inventory of affricates is more limited than German, including only postalveolar affricates [tʃ dʒ] (e.g., *chip* [tʃɪp], *jam* [dʒæm], although the alveolar affricate [ts] sometimes surfaces in loan words (e.g., Japanese *tsunami* [tsu.'na.mi] 'tidal wave'). The German labiodental affricate [pf] is novel to American English speakers.

German and English have remarkably similar syllable phonotactics. Due to the profound impact of incorporating many lexical items from Latin, French, and other languages, and the use of non-Germanic roots for forming new compounds in English, the syllable and foot structures of English and German differ most in terms of polysyllabic lexical items. But in terms of monosyllables, they are near-twins with regard to sonority sequencing restrictions on syllable margins and maximal syllable templates. For the purpose of this study, it is sufficient to

¹ For this article, which investigates the perception of naïve (pre-learner) listeners (§2.1), we do not discuss the question of [ç]-[x]-[χ] allomorphy, still an active area of research and debate in German phonology.

state that both languages allow CCVCC syllables and that both respect the Sonority Sequencing Principle (Selkirk, 1984), whereby sonority of syllable margin segments rises toward the syllable nucleus (or at least does not fall) and falls (or at least does not rise) toward the syllable margins. Specific analyses of English and German differ in the details of sonority sequencing; however, the only cluster environments employed by this study are complex codas of the (C)CV_[t] form, which is a common phonological context for occurrence of [ç x ʃ k] in German (e.g., *echt* [ɛçt] ‘genuine,’ *acht* [axt] ‘eight,’ *tauscht* [tāuʃ-t] ‘swap, 3P.SG./2P.PL.IND.PRES.,’ *fragt* [frak-t] ‘ask, 3P.SG./2P.PL.PRES.IND.’) and of [ʃ k] in English (e.g., *mashed* [mæʃ-t], *act* [ækt]). We use only VC[t] coda clusters in this study to capitalize on this phonotactic similarity between German and English.

As part of this study’s investigation of the interaction between perceptual assimilation and phonotactics in perceptual assimilation, it is necessary to consider that some segments are (not) licensed to appear in a particular position in the structure of a syllable. For this study, it is relevant that a ban against [h] in syllable codas holds for English and German (e.g., *[ti:h]), despite being permitted in some other languages, such as Turkish *tahta* [ˈtah.ta] ‘(wooden) board,’ Persian شاه *šāh* [ʃɑ:h] ‘king,’ or Arabic فواكه *fawaākih* [fa.ˈwæ:kɪh] ‘fruit, PL.’ In short, in both English and German, [h] may only occur in the pretonic simple onset position of a syllable. For a detailed account of the English distribution and proposed analysis, see Davis and Cho (2003). The participants in this study had no experience with German, and so the analogous German phonotactic restriction has no bearing on their performance in perceptual assimilation.² However, the English ban against Coda-[h] is anticipated to influence the participants’ perception in this study, in that the distribution of [h] does not transcend onset and coda positions, and so contributes to asymmetry between them in perceptual assimilation.

2. Method

2.1. Participants

Twelve native speakers of American English (nine males, three females) living in the Midwestern United States were recruited by convenience sampling in the field on the basis of their limited exposure to any foreign languages. In particular, the sample avoided those who had previous experience with German or other languages known to include [ç] or [x] in their phoneme inventories. Participant ages ranged 23–39 years ($M = 32.1$ years, $SD = 4.76$). Participants received no compensation for participation in this study.

² Nonetheless, the limited distribution of [h] in German had to be deliberately overcome by the L1 German speakers for the sake of recording the stimuli (see §2.2.2).

2.2. Stimuli

2.2.1. Phonological conditions

The perceptual assimilation experiment included balanced conditions for six German consonants [ç x h k ʃ p̄f], two adjacent vowel contexts [a ε], and three syllable positions: Simple onset, simple coda, and complex coda.³ This yielded a total of 36 combinations of these conditions: Six German consonants in six possible prosodic and adjacency environments.

2.2.2. Stimulus preparation

For each of the 36 combinations of phonological conditions, four monosyllabic nonwords were generated for a total of 144 unique item types.⁴ Whenever possible, similarity of form to English words was avoided; no item had precisely the same phonetic form as any English word.⁵ Two phonetically trained female native speakers of German each recorded three or more tokens of each item in citation form, aiming for a uniformly falling, declarative intonation.⁶ The researcher selected two tokens of each item on the basis of recording quality and consistency of intonation (as much as possible). All stimuli were equalized to a mean volume of 65 dB using a Praat script (Boersma & Weenink, 2014; Version 5.3.66). Each nonword item was instantiated by four unique token trials, two by each voice, for a total of 576 trials in the experiment. For a complete list of items, see **Table 1**.

2.3. Design

The experiment started with a training phase consisting of 12 training trials (four per syllable position), English words recorded by a female native speaker of American English. Each training phase trial was followed by an explanation of which category was the anticipated response; this included highlighting the sound in the word by means of textual enhancement of certain graphemes in the word's standard English orthography. Three experiment blocks, each focusing on a single syllable position condition, included 192 trials balanced for target consonant, adjacent

³ Most perceptual assimilation experiments control for segmental and prosodic context rather than manipulating it as a variable (e.g., variable consonant + /a/ syllable onsets only in Guion, Flege, Akahane-Yamada, & Pruitt, 2000) and do not investigate even low-level prosodic structure as a variable (e.g., syllable constituency). The work of Park & de Jong (2008, 2017) is an important counterexample to this trend; see §1.2.

⁴ Precautions were taken to avoid the German lexicon as well, but a few actual German words (e.g., [ʀɛçt]/[ʀɛxt] *Recht* 'right, law, justice') and some near homophones (e.g., [daç] vs. *Dach* [dax] 'roof') were included. As the target group for this experiment was L1 English listeners who were naïve to German, this poses no problem for the experiment.

⁵ Of the 144 items, 13 were similar enough to English words that they might be construed as accented English: [xal], [hal] vs. [haɪ] *hall*; [çes] vs. [tʃɛs] *chess*; [dɛht] vs. [dɛkt] *decked*; [ʃan] vs. [ʃɔ:n] *Sean/Shawn*; [hɛl] vs. [hɛɪ] *Hell*; [ʃɛm] vs. [ʃeɪm] *shame*, [ʃæm] *sham*; [hɛŋ] vs. [hæŋ] *hang*; [bak] vs. [bɔk] *balk*; [gak] vs. [gɔk] *gawk*; [kal] vs. [kɔɪ]/[kɑɪ] *call/cawl*; [kam] vs. [kʌm] *come*; [kɛŋ] vs. [kɪŋ] *king*.

⁶ Recordings were mono with a sampling rate of 44100 Hz.

vowel, and talker. The blocks were presented in random order, and trials were randomized within each block. At intervals of 32 trials, participants had the option to take a break.

Consonant	Phonotactic Context					
	Simple Onset		Simple Coda		Complex Coda	
	[a]	[ɛ]	[a]	[ɛ]	[a]	[ɛ]
Comparison						
[h]	hal	hɛl	gah	gɛh	baht	bɛht
	ham	hɛŋ	jah	jɛh	daht	dɛht
	han	hes	tah	tɛh	gaht	gɛht
	has	het	zah	zɛh	taht	tɛht
[k]	kal	kɛl	bak	gɛk	bakt	bɛkt
	kam	kɛm	gak	jɛk	gakt	gɛkt
	kaŋ	kɛŋ	rak	lɛk	rakt	jɛkt
	kas	kes	zak	zɛk	zakt	mɛkt
[ʃ]	ʃam	ʃɛm	baʃ	bɛʃ	baʃt	bɛʃt
	ʃan	ʃɛn	daʃ	dɛʃ	daʃt	dɛʃt
	ʃaŋ	ʃɛŋ	jaʃ	gɛʃ	jaʃt	jɛʃt
	ʃas	ʃɛt	taʃ	zɛʃ	zaʃt	zɛʃt
Novel						
[x]	xal	xɛm	gax	bɛx	baxt	bɛxt
	xam	xɛŋ	jax	gɛx	daxt	dɛxt
	xan	xɛs	rax	lɛx	gaxt	rɛxt
	xas	xɛt	zax	mɛx	raxt	zɛxt
[ç]	çal	çɛm	daç	dɛç	baçt	dɛçt
	çam	çɛn	jaç	gɛç	daçt	lɛçt
	çan	çɛs	taç	nɛç	jaçt	rɛçt
	ças	çɛt	vaç	vɛç	zaçt	zɛçt
[p̄f]	p̄fam	p̄fɛm	bap̄f	bɛp̄f	bap̄ft	bɛp̄ft
	p̄faŋ	p̄fɛŋ	dap̄f	gɛp̄f	gap̄ft	gɛp̄ft
	p̄fas	p̄fɛs	gap̄f	lɛp̄f	jap̄ft	mɛp̄ft
	p̄fat	p̄fɛt	tap̄f	tɛp̄f	vap̄ft	tɛp̄ft

Table 1: Corpus description of nonword stimuli in the current study by consonant and phonotactic context.

2.4 Procedure

Data collection in the field was conducted in private or public locations that were convenient for participants, such as a private residence ($n = 8$), a public library ($n = 3$), or a quiet coffee shop ($n = 1$), with a laptop computer.⁷ Participants first completed a five-page linguistic background questionnaire that took between five and 10 minutes (see Appendix 1). The experiment was run in OpenSesame (Mathôt, Schreij, & Theeuwes, 2012; Version 2.8.0).

Each perceptual assimilation trial consists of two tasks: Forced-choice categorization and assigning each stimulus a goodness-of-fit rating to the chosen category. Participants first saw a blank screen with one of three fixation items for 250 ms. In the fixation items, “_xx” reminded the participant to focus on the sound in initial position, “xx_” reminded the participant to focus on the sound in final position, and “xx_x” reminded the participant to focus on the sound in penultimate position (the final sound in these trials was always [t]). After 250 ms of the fixation item, the audio stimulus played. Participants heard the stimuli in stereo through a noise isolating Sennheiser HD 515 headset. They had the option to adjust the volume; during the training phase, all indicated that they could hear the stimuli clearly. Upon the end of audio playback, the screen updated to display six English categorization option boxes arranged radially around the fixation item. Participants were presented with the same six response options for every trial: “h as in *hay*,” “k as in *kite*,” “sh as in *shoe*,” “p as in *pill*,” “f as in *fun*,” and “ch as in *chew*” (see Appendix 2). These example words were selected to draw participants’ focus to the English phonemic categories of /h/, /k/, /ʃ/, /p/, /f/, and /tʃ/, respectively. The experiment recorded the location of the click for the forced-choice categorization as a set of x- and y-coordinates. After participants selected what they considered to be the most appropriate English category for each trial, the screen updated to show a seven-point scale (1 = “very bad example;” 7 = “very good example”) and a prompt to rate the item for GF to the category that had been selected by checking the box under the rating number. Once a rating was given, participants clicked an “OK” button to proceed to the next trial. All responses were given by mouse click; the task was self-paced, so trials did not time out. It took approximately 90 minutes to complete the perceptual assimilation experiment.

2.5. Data preparation

To determine which response option was chosen for each trial, the x- and y-coordinates recorded for each forced-choice categorization mouse click were sorted and labeled with the appropriate response option using a script in the R (Version 3.0.3) statistical package. Coordinates for a grid of classification boundaries were set according to the locations of the response option boxes on

⁷ Laptop hardware specifications were as follows: ASUS Notebook U56E laptop computer with an Intel Core i5-2410M 2.30 GHz processor and 6.0 GB of RAM running the Windows 7 Home Premium 64-bit operating system.

the screen (Appendix 2) with margins split evenly between categories. Responses falling outside the designated sections of the grid were not categorized by the script.

2.6. Data trimming

Fourteen mouse clicks for assigning a trial to an English category in the forced-choice task portion were positioned such that the R script could not determine what the intended category response was. These trials were trimmed from the data set before analysis. In all, of the 6912 category classification responses possible (576×12), the 14 that were trimmed for this reason represented a negligible proportion of responses (approximately 0.20%). Due to mouse click errors on the GF rating screen that failed to mark one of the checkboxes for values between 1 and 7, 13 trials did not receive a GF rating from a particular participant. These 13 trials were counted for totals of trials assigned to each response category in the forced-choice task portion, but they were excluded from calculations of GF. As there were a maximum of 576 responses per participant, and because each test item appeared four times in the experiment, these missing data points do not compromise the overall results of the present study. Overall, 6898 trials provided usable data for category choice counts, and 6885 trials provided usable data for GF ratings.

2.7. Analysis

The proportion of trials from each consonant condition (i.e., German phone) that were assigned to each forced-choice response category (i.e., English orthographic representation) was computed across participants. Similarly, the mean GF for each pairwise perceptual assimilation mapping (e.g., [h]-<h>, [ç]-<h>, [x]-<h>) was computed across participants from the available GF data. This analysis yielded two dependent variables for each pairwise mapping: Proportion of total trials by condition and mean GF for each mapping. These dependent variables were subsequently used to calculate two derivative variables: A weighted proportion (Park & de Jong, 2008) and an overlap score (Levy, 2009).

2.7.1. Weighted proportions

To integrate the forced-choice categorization frequency and GF data into a single index, we follow the approach of Park and de Jong (2008, p. 710) to derive weighted proportions using the formula shown in (1).

- (1) Weighted proportion of L2 category X in L1 category Y
 Proportion of L2 category X in L1 category Y = {probability of L2 category X is perceived as L1 category Y \times (its mean similarity score-3.5)} / Σ {probability of all L2 categories associated with L1 category Y \times (its mean similarity rating score-3.5)}.

In the formula shown in (1), the minimum threshold value for mean GF stipulated in the formula by Park and de Jong (2008)—namely, 3.5—is an arbitrary parameter set on the basis of the data in their study.

For this study, the number of response categories was six (cf. Park & de Jong's 13). Thus, a participant randomly categorizing sounds would place approximately 16% of all stimuli into each response alternative; a conservative 15% threshold is adopted for the present analysis. Furthermore, in the present data compiled from all conditions, the lowest mean value for any segment that was placed into a response category at least 15% of the time adjacent to both vowels was 3.100; however, in some conditions, that threshold is less than 3.000 (see **Table 2**).

Vowel	All Positions	By Syllable Position		
		Onset	Coda	
		Simple: _VC	Simple: CV_	Complex: CV_t
Both	3.100	3.142	3.132	3.014
[a] only	3.003	2.937	3.080	2.982
[ɛ] only	3.196	3.342	3.177	3.052

Table 2: Lowest mean GF for any mapping of 15% or more responses.

Note. On the basis of the actual mean GF minimums computed for all syllable positions combined, the mean GF threshold for the present study was set at 2.900 and used for calculation of weighted proportions using the formula of Park and de Jong (2008).

For the present analysis, the arbitrary minimum mean GF threshold was set at 2.900. This is slightly less than the actual minimum mean GF for either vowel condition and their composite when all syllable position conditions are considered, excluding responses that occur less frequently than chance level. In principle, this calculation follows the cutoff of 7% implemented by Park and de Jong (2008, p. 710) and of 7.69% by Park and de Jong (2017, p. 18). Due to this cutoff, certain position and consonant conditions (subsets of our data) exhibit slightly lower or higher minimum mean GF than those for all syllable positions, but few deviated far from this arbitrary minimum threshold, and no full position condition dipped below it. To reflect the arbitrary minimum mean GF setting for the present analysis, we employed the modified formula shown in (2) to derive weighted proportions. This low threshold component of the weighted proportion formula is applied universally, though it causes the formula in (2) to yield negative weighted proportions for some conditions. This decision affects the degree of increase of dynamic range of GF used for the formula differently between position conditions of this study, but it allows direct comparability of corresponding tables and figures across conditions.

(2) Weighted proportion of L2 category X in L1 category Y (present study)

Proportion of L2 category X in L1 category Y = $\{\text{probability of L2 category X is perceived as L1 category Y} \times (\text{its mean similarity score}-2.9)\} / \sum \{\text{probability of all L2 categories associated with L1 category Y} \times (\text{its mean similarity rating score}-2.9)\}$.

The resulting weighted proportions require some explanation for interpretation. First, as an index of category assimilation (raw) frequency and GF, weighted proportions can range negative due to certain interactions between the frequency of category selection for a phone, the mean GF for perceptual assimilations of that phone, and the minimum GF threshold (i.e., 2.9). For example, we observe a weighted proportion of -0.272 for the [ç] condition in simple codas preceded by [a] and categorized as <ch>, the lowest weighted proportion of any intersection of conditions analyzed; however, negative weighted proportions in our data represent low-frequency selections (i.e., <15% of responses for that phone) and typically have low absolute values. In the present data set, exclusion of below-chance category responses yields four noteworthy negative weighted proportions; to facilitate interpretation of these values, we provide the unweighted proportion of responses and specific mean GF for these mappings. Weighted proportions may also be depressed as compared to raw proportions as a result of a task effect from the Likert scale used to solicit subjective GF ratings. Participants do not assign uniform maximum GF to stimuli even for phones that ostensibly match their L1 categories, and for poor matches, the effect is even more pronounced. As noted by Park and de Jong (2008, p. 710), “If the mean goodness rating for a specific choice is low, it indicates that the listeners chose the answer because there were no better alternatives.” A third effect tempering such weighted proportions is the number of L1 labels selected for any particular stimulus type. If multiple labels are selected for a particular type of stimulus with any frequency, then there is a weaker connection between the stimulus phone and the L1 category; less divergent labeling indicates a stronger mapping connection.

When nonnative stimuli contain tokens of categories that align closely with the acoustic parameters of an analogous L1 category, the perceptual assimilation task normally leads to high raw frequency of mapping between the FL and L1 categories as well as consistently high GF. In other words, when two categories across languages are (nearly) equivalent, then each FL token constitutes a “very good” example of some L1 category. High frequency, paired with high GF, yields weighted proportions approaching or even exceeding 1.000. Section 3 reports the results primarily in terms of these derived weighted proportions.

2.7.2. Overlap scores

In addition to the correspondence and similarity relationships between the L2 German phones and L1 English categories measured by weighted proportions, the relationships between L2 German phones may be examined by means of the overlap scores of their mappings. Following

Levy (2009, p. 2678), overlap is defined as the frequency with which two nonnative phones perceptually assimilate to the same L1 category. By definition, every mapping of a phone to a category in perceptual assimilation has complete overlap (i.e., 1.000) with itself. If two phones are never mapped to the same L1 category, then their overlap score is 0.000. Section 3 includes overlap scores as a secondary measure of the perceptual assimilation results.

3. Results

3.1. Comparison conditions [h k ʃ p̄f]

None of the comparison conditions [h k ʃ p̄f] had a weighted proportion lower than 0.974 in any of the syllable-position conditions. This constitutes a ceiling effect with the perceptual assimilation task and confirms that the German instances of these sounds were good exemplars of corresponding English categories.

The comparison conditions with German [h k ʃ p̄f] all show categorical mappings, each to a single analogous L1 English category, and these mappings are consistent across all three syllable positions. For this reason, the remainder of the results reported here focuses on the German dorsal fricative [x ç] conditions and how their mappings vary by syllable position.

The only novel phone condition—[p̄f]—warrants specific consideration here. English lacks a labial affricate in its inventory; however, [pf] sequences do occur in morphologically derived environments in English (e.g., *helpful*, *cupful*), and so one might expect L1 English listeners to interpret [p̄f] as [p] and [f]. In the context of the perceptual assimilation task, it might then follow that in simple onset position, the “first sound” might be perceived as [p], in simple coda position, the “last sound” might be perceived as [f], and that in complex coda position before a final [t], the “second to last” sound might be perceived as [f]. However, the results show that the novel [p̄f] is categorically assimilated to English <f> in all positions (with a high GF), which fails to support the [p] + [f] sequential interpretation hypothesis. Weighted proportions show that [p̄f] assimilates categorically to <f> in simple coda (e.g., [bap̄f]: <f> 0.997, <p> -0.001) and complex coda (e.g., [bap̄ft]: <f> 0.974, <p> 0.022) positions, as well as in simple onset position (e.g., [p̄fam]: <f> 0.991, <p> 0.007). If the sequential hypothesis had been confirmed, the assimilation to <p> would have been higher in onset position. Given this result, acoustic details of these [p̄f] tokens warrant consideration.

As shown in **Table 3**, duration of the [p]-portion of [p̄f] is similar in all positions, allowing for the duration of silence of the initial closure in onset position. Duration of the [f]-portion is shorter in onsets, intermediate in complex codas preceding [t], and longest in final position, because nothing follows to close the syllable. Although shorter relative duration overall might be argued to make the [p]-portion less salient than the [f] portion in onsets and simple codas, their approximately equal duration in complex codas undermines this hypothesis. If release duration

of the [p]-portion is considered relative to duration of the [f] portion instead, then duration ratios ranging from 1:4 to 1:12 may underlie the categorical perceptual assimilation of German [pf̥] to English <f> in all syllable positions. We observed no difference in durations due to adjacent vowel context ([a] vs. [ɛ]).

Syllable position	Duration (ms)			
	[p]		[f]	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Simple onset	23	11	87	21
Simple coda	147	28	255	46
Complex coda	126	25	120	21

Table 3: Duration analysis of closure, release, and frication for [pf̥] stimuli.

Note. In simple onsets, measurement of [p] duration includes only the release portion; the preceding closure silence was not retained in the audio files. Articulations in coda position were comparable: About 100–120 ms of closure followed by an average of 20 ms of release.

3.2. Velar and palatal fricatives [x ç]

We now turn to the consonant conditions of primary interest: The dorsal fricatives [x] and [ç]. Results are presented separately for each syllable position (simple onset, simple coda, complex coda), followed by a summary of these conditions across positions. We note here an important difference between this study and Park and de Jong (2008, 2017). Their data tables display bidirectional matrices of unweighted response frequencies paired with mean GF ratings—that is, the *input* to their weighted proportion calculations—and their figures depict unweighted frequencies. We display the weighted proportions explicitly (i.e., the formula's *output*), so as to integrate both response frequency and mean GF into a single index of mapping relationship strength. As a result, our data are displayed in unidirectional tables rather than matrices—that is, each row represents all proportions for each German consonant condition and totals 1.0, but columns do not represent proportions of each L1 English response option.⁸

3.2.1. Onset position

The perceptual assimilation mappings in simple onset position according to the weighted proportions are depicted in **Figure 1**. Here it is striking that all German conditions map singularly to one English category in simple onsets, with the exception of the multiple mappings of [ç] onto three English categories. Although the mapping of [ç] to <sh> is clearly dominant,

⁸ Total values in some rows vary ± 0.001 . This is due to rounding.

the additional mappings to <h> and <ch> indicate that naïve L1 English listeners detect a difference in the acoustic signal that marks German [ç] as something other than a case of [ʃ], represented for them in English as <sh>.

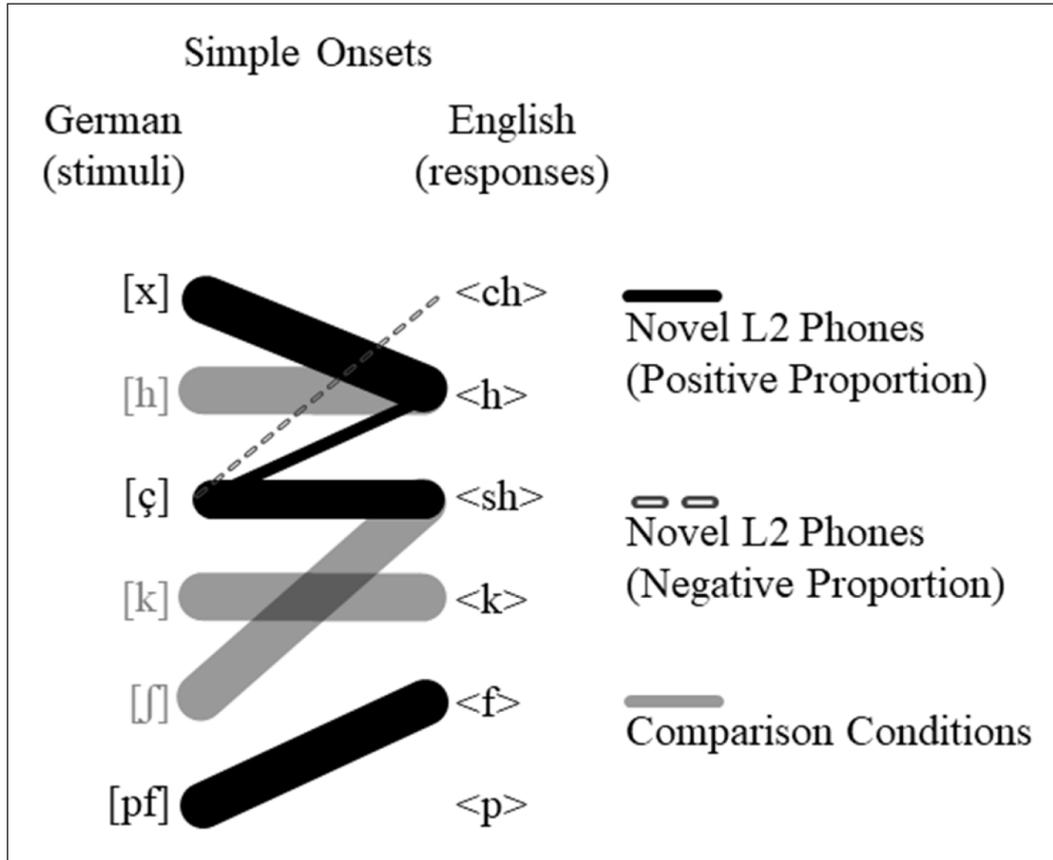


Figure 1: Proportionally weighted perceptual mapping of German phones to English orthographic categories in simple onset position based on the data from Table 4. The width of the line corresponds to the weighted proportion (i.e., proportion of instances that the English label (right) was applied to the German production (left), weighted by mean GF). Any mappings with a weighted proportion of absolute value less than 0.05 are not shown.

Table 4 displays the weighted proportions for the onset position, which integrate the frequency and GF of the mappings of all consonant conditions to all response options (Park & de Jong, 2008). The weighted proportions may be interpreted as an index of the strength of relationship between each German consonant phone condition and each English category, as represented by an orthographic response option.

In onset position, all mapping relationships are categorical (i.e., the weighted proportion is approximately 1.000), except those of German palatal fricative [ç] with English categories.

Viewed this way, [ç] in simple onsets is overwhelmingly associated with <sh>, but an interesting pattern arises between the associations with <h> and <ch>, respectively. The German phone [ç] has an above-chance positive mapping to <h> (0.274) that is partially counterbalanced in the weighted proportions by the negative association of [ç] with <ch> that remains below the $\pm 15\%$ chance level (-0.095).⁹ This notable negative weighted proportion indicates a below-chance selection frequency (0.070), further depressed by a mean GF for the mapping (2.200) that falls well below the minimum mean GF threshold (2.900) for responses that were selected at above-chance frequency.

German phone	American English category assimilations					
	<h>	<k>	<ch>	<f>	<p>	<sh>
[x]	1.000	<i>0.000</i>	<i>-0.001</i>	<i>-0.001</i>	<i>0.000</i>	<i>0.003</i>
[ç]	0.274	<i>0.000</i>	<i>-0.095</i>	<i>-0.003</i>	<i>0.000</i>	0.824
[h]	1.002	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>-0.002</i>	<i>0.000</i>
[k]	<i>0.000</i>	1.003	<i>-0.001</i>	<i>0.000</i>	<i>-0.002</i>	<i>0.000</i>
[pf]	<i>0.000</i>	<i>0.002</i>	<i>0.001</i>	0.991	<i>0.007</i>	<i>-0.002</i>
[ʃ]	<i>0.001</i>	<i>0.000</i>	<i>0.000</i>	<i>0.006</i>	<i>0.000</i>	0.994

Table 4: Weighted proportions assimilated to each response option: Simple onset position. *Note.* Conservative chance categorization threshold was 15%. Lowest mean GF for any segment of 15% or more responses was 3.142 in this position; minimum threshold of 2.900 was used for all conditions. The highest value in each row is bold; absolute values less than 0.05 are italic.

The weighted proportions reveal a markedly weaker relationship in simple onsets between English <h> and German [ç] than between English <h> and German [x]. The relatively weak strength of association between German [ç] and English <h> is primarily compensated for by the robust mapping of [ç] to English <sh> in this position, which is nearly as strong a relationship as the categorical mapping of the comparison condition German [ʃ] to English <sh>.

The relationships between the German phones of the stimuli may be examined directly by means of their overlap scores (Levy, 2009), shown in **Table 5**. The uniform overlap scores of the diagonal (i.e., 1.000) indicate that each German consonant's mapping is completely consistent with itself. Overlap scores that approach 1.000 approach the level of identity, whereas those approaching 0.000 approach complete differentiation of the conditions by means of their dissimilar mappings. Consistent with the weighted proportions, the degree of overlap between [x] and [h]

⁹ If the condition-specific minimum mean GF threshold of 3.142 were retained for calculation of Park and de Jong's (2008) formula rather than the arbitrary setting of 2.900 used for Table 1, then the resulting weighted proportion for the [ç]-to-<ch> mapping would be -24.0% , an absolute value well above chance level.

is nearly complete in simple onset position (0.995) and [ç] has a much lower overlap score than [h]. Although [ç] has a strong association with <sh> in terms of weighted proportions, this relationship is not mutual: [ʃ] is categorically mapped to <sh> with no alternative mapping candidates. As [ʃ] does not also map to <h> or <ch> along with [ç], the overlap of [ç] and [ʃ] scores low. Also, in contrast to [x], overlap between [ç] and [h] is markedly lower (0.587). The similarly low overlap between [ç] and [x] (0.594) indicates that these phones are frequently mapped to the same categories but almost as frequently mapped to different categories—thus mappings of [ç] frequently diverge from the acoustic space occupied by [h] and [x] toward the space occupied by [ʃ], where overlap of [ç] and [ʃ] scores 0.364. Finally, the negligible overlap scores of [k] and [pf̩] with any other category indicate no interaction in perception of these conditions with any other conditions in simple onset position.

German phone	German phone					
	[x]	[h]	[ç]	[k]	[ʃ]	[pf̩]
[x]	1.000					
[h]	0.995	1.000				
[ç]	0.594	0.587	1.000			
[k]	0.003	0.003	0.008	1.000		
[ʃ]	0.016	0.008	0.364	0.008	1.000	
[pf̩]	0.008	0.003	0.018	0.010	0.013	1.000

Table 5: Overlap scores for all pairs of phones presented in stimuli: simple onset position.

3.2.2. Coda positions

The remaining position conditions have in common that they are syllable codas. In German phonotactics, complex codas ending in final [t] are commonplace, and so simple codas and final C[t] clusters are also considered. English has a phonotactic ban against the occurrence of [h] in syllable codas (simple or complex), which is a crucial difference from onsets (see §1.3).

3.2.2.1. Simple codas

Figure 2 depicts weighted proportions from **Table 6**. Although strength of association between [x] and <h> is strong, in this position, weak mappings of [x] also to <k> and <f> mitigate this strength. This effect is slight compared to the multiply distributed mappings of [ç] to phonotactically licit <sh> (strong), to illicit <h> (strong), to identity-ambiguous <ch> (strongly negative), and to phonotactically licit <k> (weak). Note that [x] and [ç], both fricatives, show minor mappings to <k>, representing the stop [k], in simple codas.

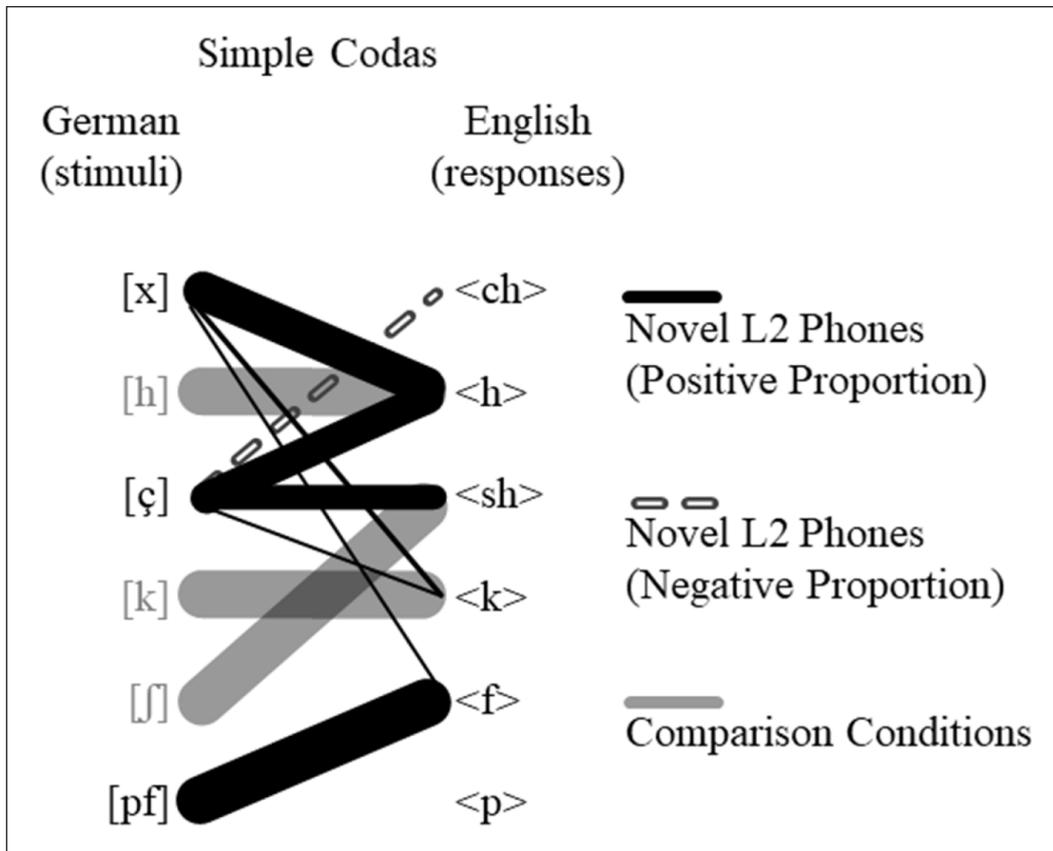


Figure 2: Proportionally weighted perceptual mapping of German phones to English orthographic categories in simple coda position based on the data from Table 6. The width of the line corresponds to the weighted proportion (i.e., proportion of instances that the English label (right) was applied to the German production (left), weighted by mean GF). Any mappings with a weighted proportion of absolute value less than 0.05 are not shown.

The weighted proportions for simple coda position are displayed in Table 6. In this position, all comparison conditions [h k p̄f ʃ] categorically map to English <h>, <k>, <f>, and <sh>, respectively. This is despite the phonotactic ban against [h] in syllable codas in English, the listeners' L1.¹⁰

Neither German [x] nor [ç] exhibit categorical mappings in simple codas. Although <h> is the clear modal response for [x] (0.836), the strength of this association is mitigated by a secondary mapping to <k> (representing [k]) that approaches chance in strength (0.120) as well as a minor tertiary mapping to <f> (representing [f]; 0.061), both of which have the advantage over [h] of being phonotactically licit in English codas.

¹⁰ The ban against [h] in codas also holds for German, but as these listeners remained naïve at the time of testing, we presume that German phonotactic principles should be irrelevant to the listeners' performance on this task.

German phone	American English category assimilations					
	<h>	<k>	<ch>	<f>	<p>	<sh>
[x]	0.836	0.120	-0.009	0.061	-0.004	-0.004
[ç]	0.660	0.091	-0.253	-0.007	0.000	0.510
[h]	1.001	-0.001	0.000	0.000	0.000	0.000
[k]	0.002	0.998	0.000	0.000	0.000	0.000
[p̥f]	0.002	0.000	0.000	0.997	-0.001	0.001
[ʃ]	0.001	0.002	-0.001	0.000	0.000	0.998

Table 6: Weighted proportions assimilated to each response option: Simple coda position.
Note. Conservative chance categorization threshold was 15%. Lowest mean GF for any segment of 15% or more responses was 3.132 in this position; minimum threshold of 2.900 was used for all conditions. The highest value in each row is bold; absolute values less than 0.05 are italic.

The variety of mappings of German [ç] are even more diverse. The modal mapping of [ç] to phonotactically illicit <h> is relatively weaker (0.660), while its secondary mapping to <sh> is a near competitor (0.510), which already seems more a bimodal response pattern. This is complicated by a strong tertiary, negative mapping to <ch>, which indicates below-chance selection frequency (0.089), further depressed by a mean GF (2.210) well below the minimum mean GF threshold (2.900) for responses that were selected at above-chance frequency. Taken together, this denotes a poorly fitting association of German [ç] with English <ch> in simple codas (-0.253). There is an additional fourth association of [ç] with <k> approaching chance level (0.091). Note that the secondary association (<sh>) and the weakest (<k>) are clearly phonotactically licit in English codas. Orthographic <ch> may represent a licit category (e.g., [tʃ], [k]) or a dialectally marked, if not illicit, category (e.g., Scottish English [x]).

In terms of overlap scores, displayed in **Table 7**, only the German [p̥f] condition remains completely aloof from the other categories in simple codas. High overlap between [x] and [ç] (0.827) is driven primarily by their common overlap with [h] in this position ([x]-[h] = 0.801; [ç]-[h] = 0.702), although both also show minor overlap with [k] ([x]-[k] = 0.102; [ç]-[k] = 0.047). Differentiated perception between [x] and [ç] finds support in this position from the overlap between [ç] and [ʃ] (0.190) that is not paralleled by [x] versus [ʃ] (0.024).

There are two key observations to draw from the data in simple coda position. First, the strength of associations between German dorsal fricatives [x ç] and English <h>, as well as the strength of association between [ç] and <sh>, is mitigated by distribution of responses across more English categories. Second, despite the fricative articulation of [x ç], in this position they both have weak associations with English <k>, which represents the stop [k] that is phonotactically licit in syllable codas, where [h] is not.

German phone	German phone					
	[x]	[h]	[ç]	[k]	[ʃ]	[pf]
[x]	1.000					
[h]	0.801	1.000				
[ç]	0.827	0.702	1.000			
[k]	0.102	0.013	0.047	1.000		
[ʃ]	0.024	0.005	0.190	0.005	1.000	
[pf]	0.042	0.013	0.023	0.008	0.005	1.000

Table 7: Overlap scores for all pairs of phones presented in stimuli: Simple coda position.

3.2.2.2. Complex codas

This study represents complex codas with one possible coda structure: C[t] clusters. The diffusion of mappings in complex coda position and the strong relationships between German [x] and [ç] and English <k> are depicted in **Figure 3**. In this position, there is no significant mapping between [ç] and <h>. Both [x] and [ç] show a near-even split between two major mappings: <h>/<k> and <sh>/<k>, respectively. In addition, [x] has a weak positive association with the ambiguous English <ch>, whereas [ç] has weak negative associations (infrequent and poor fit) with <ch> and <p>. Striking in this depiction are the approximately even mappings of both [x] and [ç] to <k> in this position.

Table 8 displays the weighted proportions for complex codas. A more diffuse response pattern in complex coda position as compared to simple onset or simple coda positions is evidenced by the greater number of cells in **Table 8** that have extremely low or negative non-zero values and by the more varied modal response patterns, although the comparison conditions [h k pf ʃ] still exhibit categorical mappings to single English response options <h>, <k>, <f>, and <sh>, respectively. German [x] shows clear mapping to both English <h> (0.495; primary) and <k> (0.488; secondary), as well as weak mapping to English <ch> (0.071 < 15% chance threshold). Similarly, the dominant mapping of German [ç] to English <sh> (0.471) is contested by a comparably strong mapping to English <k> (0.461), as well as two negative weighted proportional mappings to English <ch> (-0.029) and English <p> (-0.060). These both represent below-chance selection frequency (<ch> = 0.073 ; <p> = 0.029), further depressed by mean GF ratings (<ch> = 2.750; <p> = 2.090) that fall below the minimum mean GF threshold (2.900) for responses that were selected at above-chance frequency.¹¹

¹¹ For discussion of <p> as a specific perceptual assimilation phenomenon rather than mere variation, cf. Park & de Jong (2017).

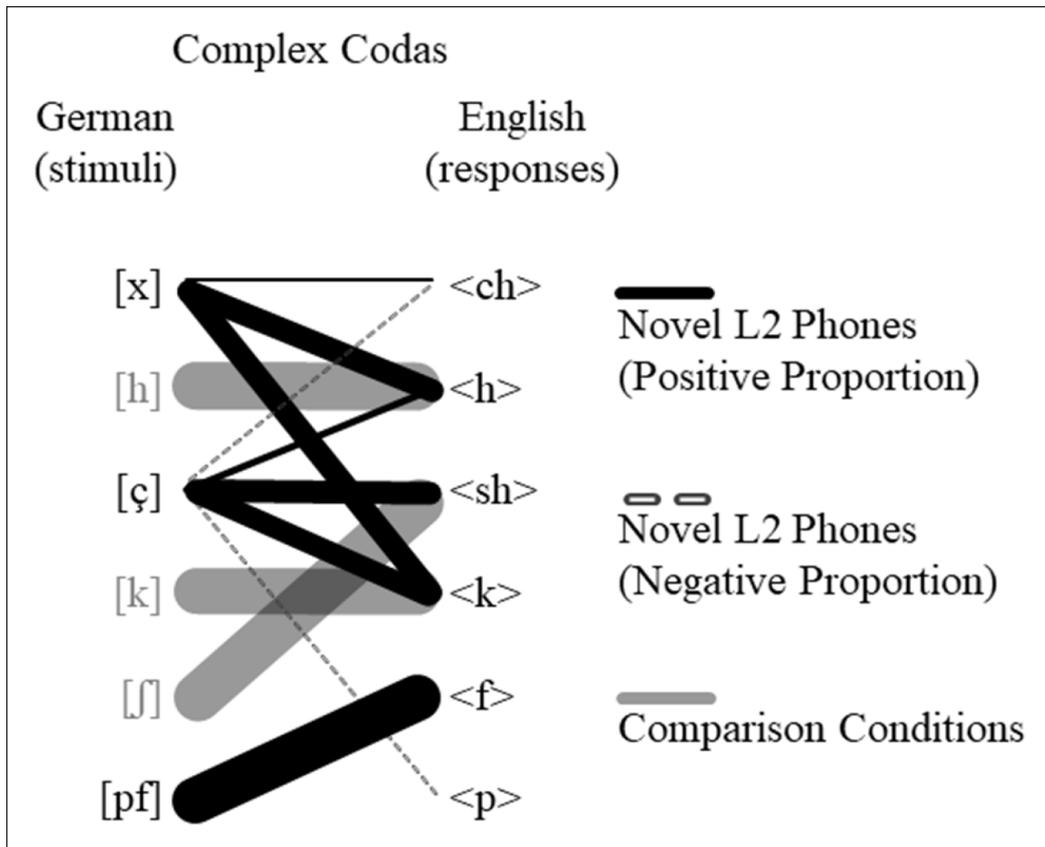


Figure 3: Proportionally weighted perceptual mapping of German phones to English orthographic categories in complex coda position based on the data from Table 8. The width of the line corresponds to the weighted proportion (i.e., proportion of instances that the English label (right) was applied to the German production (left), weighted by mean GF). Any mappings with a weighted proportion of absolute value less than 0.05 are not shown.

German phone	American English category assimilations					
	<h>	<k>	<ch>	<f>	<p>	<sh>
[x]	0.495	0.488	0.071	<i>-0.005</i>	<i>-0.030</i>	<i>-0.020</i>
[ç]	0.161	0.461	<i>-0.029</i>	<i>-0.004</i>	<i>-0.060</i>	0.471
[h]	1.012	<i>0.003</i>	<i>-0.005</i>	<i>0.013</i>	<i>-0.015</i>	<i>-0.007</i>
[k]	<i>0.012</i>	0.979	<i>0.005</i>	<i>0.000</i>	<i>0.004</i>	<i>0.000</i>
[pf]	<i>0.004</i>	<i>0.000</i>	<i>0.000</i>	0.974	<i>0.022</i>	<i>0.000</i>
[ʃ]	<i>0.002</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	0.998

Table 8: Weighted proportions assimilated to each response option: complex coda position. *Note.* Conservative chance categorization threshold was 15%. Lowest mean GF for any segment of 15% or more responses was 3.014 in this position; minimum threshold of 2.900 was used for all conditions. The highest value in each row is bold; absolute values less than 0.05 are italic.

In terms of relationship strength, the weighted proportions for complex coda position, in which the listening target precedes a final [t], show both wider variety of mappings that warrant consideration and a clear tendency to interpret German [x] and [ç] as some kind of English <k> (representing [k]), despite the mismatch between the fricative articulation of the German consonants and the stop articulation of the English category to which they are mapped.

The overlap scores displayed in **Table 9** show that only German [pf̂] has little overlap with the other German consonant conditions; the greater number of overlap scores notably above zero indicates here that more of the stimuli conditions were mapped with some frequency to more of the response categories. In this position, overlap between [x] and [h] is high (0.728), whereas overlap between [ç] and [h] is relatively lower (0.597). Nonetheless, high overlap between [x] and [ç] (0.835) is maintained by virtue of their mutual overlap with [k] ([x]-[k] = 0.275; [ç]-[k] = 0.243). Meanwhile, perceptual distinction between [x] and [ç] is maintained in this position by the overlap of [ç] and [ʃ] (0.199), which is not paralleled by [x] or [h]. Finally, in complex coda position (preceding [t]), [k] exhibits noteworthy overlap with all other phones except [ʃ] and [pf̂], ranging from [h] (0.070) to [x] (0.275).

German phone	German phone					
	[x]	[h]	[ç]	[k]	[ʃ]	[pf̂]
[x]	1.000					
[h]	0.728	1.000				
[ç]	0.835	0.597	1.000			
[k]	0.275	0.070	0.243	1.000		
[ʃ]	0.045	0.021	0.199	0.018	1.000	
[pf̂]	0.037	0.037	0.045	0.013	0.005	1.000

Table 9: Overlap scores for all pairs of phones presented in stimuli: Complex coda position.

Overall, the results for the complex coda position before a final [t] show that naïve L1 English listeners readily associate the German consonants [h x ç ʃ] with [k], which seems to result in notably more confusion between them, although this effect is small for [h] and [ʃ].

3.3. Summary of results

The focused results by syllable position in the preceding sections highlight a variety of local perceptual assimilation mapping relationships between phones that can be further examined across positions to better understand the tendencies of each German phone that remain consistent in all positions. To this end, the results for German [x] and [ç] from all positions are

reorganized and collated for side-by-side comparison in this section. The mappings of all the German consonant conditions across all positions are depicted in **Figure 4**, which reproduces **Figures 1, 2, and 3** side-by-side for comparison.

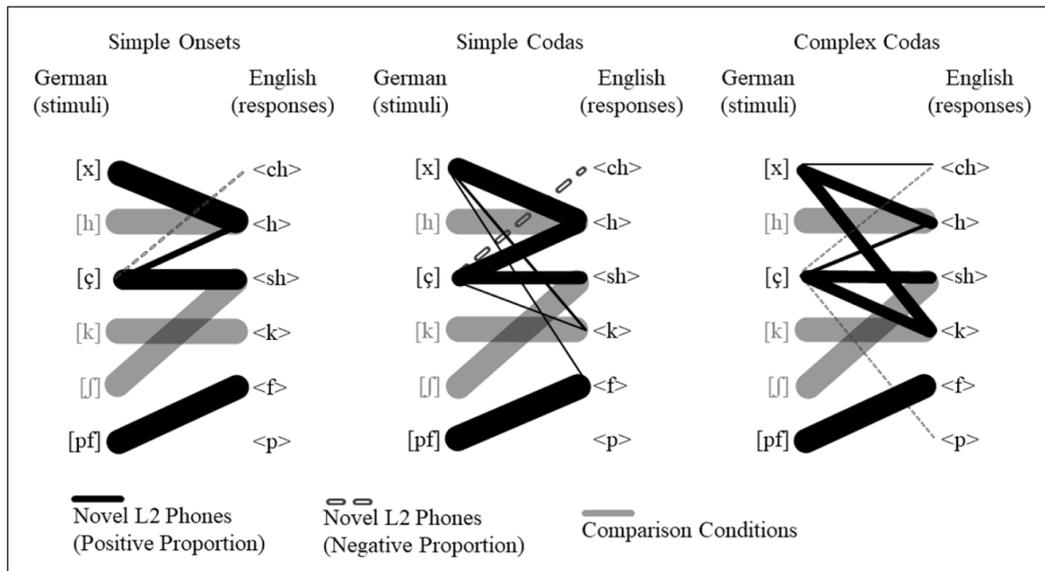


Figure 4: Proportionally weighted perceptual mapping of German phones to English orthographic categories in each tested syllable position, following Park and de Jong (2008, p. 710). The width of the line corresponds to the weighted proportion (i.e., proportion of instances that the English label (right) was applied to the German production (left), weighted by mean GF). Any mappings with a weighted proportion of absolute value less than 0.05 are not shown.

Visual inspection of the diagrams in **Figure 4** highlights the two effects observed in the data. The first effect manifests between onsets and codas. As compared to simple onset position (left), both coda positions (center and right) exhibit both a wider variety of mappings for German [x] and [ç], represented by lines connecting each to more English orthographic categories, including the stop <k>, which draws no mappings from these fricatives in simple onsets. This finding with a different language pairing (L1 English-L2 German vs. L1 Korean-L2 English; cf. Park & de Jong, 2008, 2017) provides independent evidence for differential perceptual mapping patterns between onsets and codas across languages and additionally sheds light on the interaction of segmental manners of articulation with low-level prosodic complexity (i.e., subsyllabic constituency) in perception.

The second effect is shown by comparison of the two coda types. Between simple codas (center) and complex codas (right), the number of mappings (i.e., lines) that arise for [x] and [ç] is the same. However, for both German fricatives [x ç], the increased strength of relationship with English <k> in complex coda position (right), at the expense of their relationship strengths

to <h> (and also of [ç] to <sh>), has the result that no single mapping to an English category emerges as clearly dominant for either [x] or [ç]. Furthermore, only in complex coda position, where the consonant cluster includes an adjacent final stop [t], does <k>, representing stop [k], arise as a serious competitor for mapping these German fricatives.

These two observations—greater diversity of mappings in codas generally and increased mapping of fricatives [x] and [ç] to <k> despite the difference in manner of articulation—illustrate major asymmetries in perceptual assimilation according to prosodic environment (i.e., syllable onsets vs. codas) and phonotactic environment (i.e., singleton consonants vs. consonant clusters). To enable more detailed direct comparisons across positions, **Tables 10** and **11** compile all weighted proportions for [x] and [ç], respectively, from **Tables 4, 6, and 8**.

Weighted Proportion		Response options: American English categories					
		<h>	<k>	<ch>	<f>	<p>	<sh>
Position	Simple Onset	1.000	<i>0.000</i>	<i>-0.001</i>	<i>-0.001</i>	<i>0.000</i>	<i>0.003</i>
	Simple Coda	0.836	0.120	<i>-0.009</i>	0.061	<i>-0.004</i>	<i>-0.004</i>
	Complex Coda	0.495	0.488	0.071	<i>-0.005</i>	<i>-0.030</i>	<i>-0.020</i>

Table 10: [x]: Weighted proportions for each response option by position.

Note. Conservative chance categorization threshold was 15%. Minimum threshold of 2.900 was used for all conditions. The highest value in each row is bold; absolute values less than 0.05 are italic.

Weighted Proportion		Response options: American English categories					
		<h>	<k>	<ch>	<f>	<p>	<sh>
Position	Simple Onset	0.274	<i>0.000</i>	<i>-0.095</i>	<i>-0.003</i>	<i>0.000</i>	0.824
	Simple Coda	0.660	0.091	<i>-0.253</i>	<i>-0.007</i>	<i>0.000</i>	0.510
	Complex Coda	0.161	0.461	<i>-0.029</i>	<i>-0.004</i>	<i>-0.060</i>	0.471

Table 11: [ç]: Weighted proportions for each response option by position.

Note. Conservative chance categorization threshold was 15%. Minimum threshold of 2.900 was used for all conditions. The highest value in each row is bold; absolute values less than 0.05 are italic.

The data for German [x] in **Table 10** show two key effects of position on perceptual assimilation mappings by naïve L1 American English listeners. In simple onsets, German [x] is consistently assimilated to <h> (representing [h]), with no competition from any other potential mapping. In contrast, in both coda positions, where [h] is phonotactically illicit in English, the mapping of

[x] to <h> sees competition from other mappings to <k>, <ch>, and <f>, all of which may represent categories that are phonotactically licit in English codas, as long as <ch> for [tʃ] (e.g., *cheese*) or the variant [x] (e.g., Scottish *loch*) are considered among viable English candidates. The second effect arises in the comparison of simple codas with complex codas. In complex codas only, [x] to <k> rises to dominance, preferred to <h> despite the acoustic-spectral similarity and shared fricative manner of articulation (Stevens, 1960). In complex codas preceding the stop [t], this fricative is more readily mapped to a stop with the same place of articulation as the German [x].

The weighted proportions for German [ç] in **Table 11** show more complex perceptual assimilation patterns than [x] in all positions. In contrast to [x], German [ç] has no categorical mapping in any position. In simple onsets, the German palatal fricative [ç], which does not precisely map any place of articulation familiar to the English phonemic inventory, primarily maps to English <sh> (representing spectrally similar [ʃ]; Stevens, 1960). However, this mapping faces notable competition from two additional mappings that counterbalance each other—namely, a secondary mapping to <h> and a negative tertiary mapping to <ch>. As with [x], in both coda positions, these three mappings of German [ç] see competition from the mapping of [ç] to <k>; however, mapping to <sh>, which is phonotactically licit in English codas, remains the primary mapping in all positions. Between simple and complex codas the second effect arises again: When preceding the stop [t] in a complex coda, the mapping of [ç] to <k> rises in prominence, nearly the equal of the primary mapping to <sh>, despite the difference in manner of articulation between [ç] and [k] and their different places of articulation.

Examination of overlap scores in review illuminates additional relationships between the German phones of interest. German [p̄f] consistently had almost exclusive overlap with itself, which indicates a lack of interaction with the other phones presented in the experiment. For that reason, [p̄f] is excluded from discussion here. **Figure 5** visualizes the overlap score relationships between all other German consonant conditions by syllable position to allow for side-by-side comparison, reproducing selected data from **Tables 5, 7, and 9**.

In all three syllable positions, overlap scores between [ç] and [ʃ] are above chance level (> 16%), particularly in simple onset position (36.4%). In contrast, overlap of [x] and [ʃ] never approaches chance level in any position. German [x] and [ç] show their least overlap in simple onset position (59.4%), where both overlap of [x] and [h] (99.5%) and overlap of [ç] and [ʃ] (36.4%) are highest—that is, German [x] and [ç] exhibit mappings least similar to each other in simple onset position. In both coda positions, overlap of [x] and [ç] is higher than any other overlap score (> 80%), although closely rivalled by the overlap of [x] and [h] in simple coda position (80.1%); this suggests a prominent three-way overlap pattern including [x]-[ç] (simple coda: 82.7%, complex coda: 83.5%), [x]-[h] (simple coda: 80.1%, complex coda: 72.8%), and, secondarily, also [ç]-[h] (simple coda: 70.2%, complex coda: 59.7%).

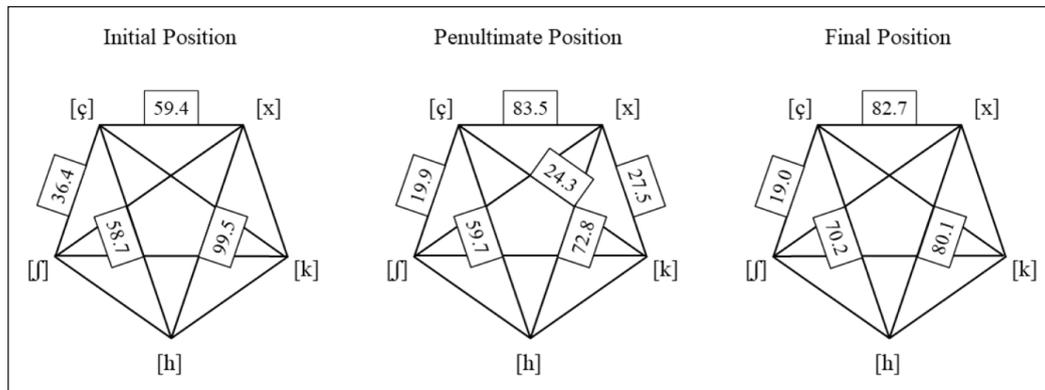


Figure 5: Overlap scores (in %) are drawn from Tables 5, 7, and 9, with the exception of $[\widehat{pf}]$, which exhibits no significant overlap with any other target phone in the stimuli. Lines connecting two phones are labelled by the overlap score for that pair of phones. Chance-level (random) overlap is approximately 16% for six stimuli types; overlap scores below 15% are omitted.

Focusing on above-chance overlap scores beyond those discussed thus far (including the overlap of $[\ç]$ and $[\ʃ]$), two patterns emerge: One global and one local to the complex coda position. In all positions, both $[x]$ and $[\ç]$ have overlap scores with $[h]$ above 50%. In complex coda position only, both dorsal fricatives also have above-chance overlap scores with $[k]$ ($[x]$ - $[k]$: 27.5%, $[\ç]$ - $[k]$: 24.3%). The simple coda position also exhibits overlap scores clearly above zero with $[k]$, but none reach chance level. Indeed, all overlap scores with $[k]$ in complex coda position are higher than their corresponding scores in either simple onset or simple coda, including a spike of nearly 7% between $[k]$ and $[h]$ (which is banned from English codas), but only overlap of $[k]$ with $[x]$ (27.5%) and $[\ç]$ (24.3%) exceed the chance-level threshold. This local pattern suggests that the complex coda trials' CV_[t] frame is responsible for the marked increase in mapping fricatives to the stop $[k]$ by causing listeners to interpret the novel dorsal fricatives $[x]$ and $[\ç]$ in this environment as sharing the manner of articulation of the adjacent coda stop $[t]$. This indicates that the local phonotactic context of the German dorsal fricative induces a stronger mapping to the English stop $[k]$ in the perceptual assimilation of naïve L1 English listeners *in addition to* the different mappings between $[\ç]$ and $[\ʃ]$, $[x]$ and $[\ç]$, and the German dorsal fricatives and $[h]$.

4. Discussion

This study set out to examine how American English listeners (without any knowledge of German) perceive novel German sounds. We asked whether and how syllable position, a low-level prosodic domain, and phonotactic legality influence the perceptual assimilation patterns for these sounds. We found that syllable position indeed affects perceptual assimilation of German dorsal fricatives by naïve L1 English listeners—that is, perceptual assimilation of these *novel*

sounds is modulated by syllable position and L1 phonotactic biases. For the sounds examined that were familiar to the phonemic inventory of the listeners' L1 English ([h k ʃ]), perceptual assimilation mappings were not affected by syllable position (i.e., onset vs. coda) nor by illicit phonotactic position (i.e., V[ht] complex codas), *despite* the phonotactic ban against [h] in coda positions in English (Davis & Cho, 2003). Note that the L1 phonotactic ban on [h] in codas may help to explain why phonotactically legal [k] gets a boost in all codas and not just next to final [t]. However, if predicting purely on the basis of L1 phonotactic biases, we would not expect naïve L1 English listeners to persist in perceiving “illegal” [h] in codas at all, which they do. Neither was perceptual assimilation of the novel affricate [pʃ] affected by syllable position. Instead, it appears that the affricate was perceived as the L1-familiar [f] without regard for its initial labial closure, the relatively short release of which may not be salient enough for L1 English listeners. In contrast to the familiar sounds (including infelicitous “[f]”), for the two novel phones [x] and [ç], we found that perceptual assimilation was modulated by both syllable position (i.e., onset vs. coda) and by phonotactic adjacency context (i.e., VC[t] complex coda). Overall, each sound exhibited a different mapping pattern in each of the three positions tested: Simple onset, simple coda, and complex onset. In contrast to the onset position, in both coda positions the novel dorsal fricatives [x ç] exhibit more diffuse (and thus individually weaker) mapping patterns and a local tendency to map to <k> (representing [k]), a segment that is phonotactically licit in English codas. Within codas, the phonotactic difference between simple and complex is reflected in [t]-final codas by stronger mappings of both [x] and [ç] to <k> *despite* the difference in manner of articulation (i.e., fricative vs. stop), which yields yet more diffusion in mapping—that is, *no* mapping is clearly dominant for either of these novel sounds in complex codas. In summary, for naïve L1 English listeners, perceptual assimilation of the German dorsal fricatives [ç x] differs markedly between onsets and codas in terms of number and strengths of relationship, and between simple codas and complex codas, where adjacency to the stop [t] reveals a bias to perceive these novel fricatives as a phonotactically licit (nearly) homorganic stop [k] rather than the locally illicit fricative [h].

The assimilation patterns for [x] and [ç] reveal that they are both categorized as some kind of English <h>; however, unlike German [h], they were not perceived exclusively as <h>. The novel German phones [x] and [ç] each have different mapping patterns to English categories other than <h>, including differences in GF as examples of the English categories onto which they are mapped. These facts indicate that, despite the prevalence of their mappings to English <h>, [x] and [ç] are perceived by naïve L1 English listeners as somehow unlike each other and not quite the same as [h]. On the basis of weighted proportions and overlap scores, relationships between [x] or [ç] and [h] find competition from their relationships with [k] in codas. Furthermore, with the exception of [x] in simple onset position, neither dorsal fricative ever shows a clearly categorical relationship with any of the English category response options that were available in this study.

It is clear that these sounds represent a significant challenge for naïve listeners, and thus, by extension, for phonological acquisition by L1 English speakers at the earliest stages of L2 German exposure. The dorsal fricatives are in complementary distribution in German, but early L2 learners receive only limited exposure to them. These two sounds are orthographically associated with <ch>, but they are acoustically different, with different perceptual assimilation patterns in every syllable position investigated here. Unless given ample opportunity to associate the acoustic signal with the correlating orthographic signal cross-modally (e.g., listening to many occurrences of these sounds while simultaneously reading a text that displays their conditioning contexts), learners typically receive little positive evidence upon which to posit such a distributional relationship between them. Moreover, these sounds do not exist in isolation: They are both confusable with other sounds present in both L1 and L2—namely, [h k ʃ]. The likelihood of these confusions in perception also varies by prosodic and phonotactic context. Thus the learning challenge differs for onsets, where [x] maps categorically to [h] while [ç] maps partly to [h] and partly to [ʃ], as opposed to codas, where English [ʃ] becomes a stronger candidate for German [ç] and [k] becomes a possible mapping for both [x] and [ç], especially in clusters. Under such circumstances, it is unsurprising that L2 German learners would have difficulty with minimal pairs such as German *Nacht* [naxt] ‘night’ and *Fracht* [fraxt] ‘freight’ versus *nackt* [nakt] ‘naked’ and *fragt* [frak-t] ‘ask, 3P.SG.PRES.IND./2P.PL.PRES.IND.’ or *stechen* [ʃtɛ.çŋ] ‘sting, INF.’ versus *stecken* [ʃtɛ.kŋ] ‘put, INF.’ Although such minimal pairs are readily disambiguated by lexical, semantic, and syntactic context by proficient language users, beginning learners have yet to develop robust online representations and processing strategies for L2 speech. As long as speech processing relies heavily on the perception and recognition of individual sounds, learners’ capacity for parsing running speech, recognizing familiar and new vocabulary—and for matching these to L2-appropriate orthographic and lexical representations—is adversely affected.

These findings with regard to perceptual assimilation of German [x] and [ç] by naïve L1 English listeners supports a broader claim in L2 phonology that L2 learners, especially beginners, rely more on subphonemic phonetic detail than native speakers or advanced L2 learners. This then makes perception of nonnative phones sensitive to specific positions and contexts (Best & Tyler, 2007; Flege, 1995; Hallé, Best, & Levitt, 1999; Lively, Logan, & Pisoni, 1993; Sheldon & Strange, 1982). Beyond the general tendency for both of these German phones to be perceived as a type of <h> in any position, this study has demonstrated that their perceptual assimilation mappings differ extensively from each other according to both subsyllabic structure (i.e., syllable onsets vs. codas) and phonotactic environment (i.e., singleton consonants vs. consonant clusters). The German fricatives [x] and [ç] manifest different patterns of mapping on the basis of syllable position, such as the tendency for [ç] to map more strongly to <sh>/[ʃ] in simple onset position than in codas and for both [x] and [ç] to map more strongly to <k>/[k] in codas than in onsets, especially under the influence of an adjacent [t] in complex codas. In contrast to novel phones, phones associated with familiar L1 phoneme categories /h k ʃ/, and the novel phone

[p̥f] *interpreted* as the familiar L1 category /f/, exhibit a robust insensitivity to position, which suggests phonemic processing rather than phonetic processing. This crucial difference suggests that phonemic processing enables—or entails—normalization of acoustic differences across multiple positions and contexts, consistent with empirical investigations and theoretical models of L2 phonological acquisition (e.g., Levy & Strange, 2008; Strange, 2011).

We show that German [x] and [ç] do not cleanly perceptually assimilate to any single English category, with the local exception that [x] is almost universally taken as a kind of <h> in simple onset position, albeit not as good an example of it as German [h]. These results would be challenging to interpret in terms of the classic PAM as being purely of the single-category (SC Type), of the category-goodness difference (CG Type), or of the two-category (TC type) assimilation types (Best, 1995, pp. 194–195); however, the assimilation patterns we observe for [x] and [ç] evidence a perceptual assimilation profile potentially suitable for investigation with the *phonological overlap method*, a recent refinement of PAM for novel phones for which there is perceived phonological similarity with one or more L1 categories (Faris, Best, & Tyler, 2018). Under the phonological overlap method, [ç] seems a prime candidate to be analysed as a *clustered* assimilation in all syllable positions investigated here. In contrast, [x] might qualify as a *focalised* assimilation in the simple onset position but a clustered assimilation in simple and complex codas. The perceptual assimilation patterns we observed for these two phones overlap as <h> in all syllable positions at chance level or higher, which fits the *partial overlap* profile developed for UU and UC assimilation types (Faris, Best, & Tyler, 2018, pp. 3–4). It remains to be investigated by future research whether the German dorsal fricatives [ç] and [x] comprise a UU, UC, or another assimilation type for L1 English speakers.

The question of how naïve L1 English speakers perceive the German dorsal fricatives is additionally complicated by the different perceptual assimilation patterns of the two phones in each of the three positions tested. Following the findings of Daidone et al. (2015, 2019), we propose that mapping analyses such as weighted proportions (Park & de Jong, 2008, 2017), which quantitatively integrate GF ratings in a way that PAM does not yet (Best, 1995; Faris, Best, & Tyler, 2018), should be considered in conjunction with overlap scores (Levy, 2009), which allow perception of FL/L2 sounds to be analyzed in relation to each other and not just in relation to L1 sounds. These rubrics also avoid the variable discriminability rate thresholds used to quantify descriptors such as “poor,” “good,” or “very good” that hinder clarity of the canonical assimilation type taxonomy proposed in the PAM, a problem still acknowledged within recent investigations in the PAM framework (Best, 1995; Faris, Best, & Tyler, 2018).

This study confirms the findings of numerous other studies showing that syllable position and phonotactic adjacency influence perception of speech sounds (e.g., Kilpatrick et al., 2018; Kilpatrick et al., 2019; Park & de Jong, 2017; Winters, 2001), but this opens the door to the greater question of *why* these factors exert such an influence on speech perception. In terms

of tonal perception (i.e., suprasegmentals), House (1996) offers what may be a helpful clue for answering this question. His model proposes that the *amount* of novel acoustic information encountered simultaneously differs at different points in the syllable, and this limits how listeners attend to and perceive pitch information relevant to phonological tone. Early in the syllable (i.e., onset and early nucleus), tonal information occurs along with a great deal of novel vowel formant information, and so tone is perceived as level. Later in the syllable (i.e., late nucleus and early coda), when the formants have typically established a steadier vowel state, tonal information effectively has less competition, and so pitch contours are perceived and recognized more felicitously. The relative attention afforded “novel” information over familiar information is also documented in the broader perception literature, which posits a “novel popout” effect, whereby reaction times may decrease (i.e., speed up) for novel stimulus types as they receive more attentional focus (Christie & Klein, 1996; Johnston & Schwarting, 1996, 1997). Further study is needed to address the question of differential perception to consonants in the onset or coda of a syllable. Within syllable margins, further investigations of phonetic masking and other adjacency and sequence effects in audition with regard to phonotactics are also needed.

Our main finding that perception of novel sounds is modulated by their syllable position has potential implications for phonological learning in a second language. If we take the naïve listeners as representing the initial stage of L2 phonological acquisition, our findings that syllable position matters for sound categorization are potentially important to refine our understanding of how learning unfolds, such as which confusions will be more likely, or in which words certain sounds are more difficult than others to distinguish and pronounce; however, we do not answer everything. Our results highlight the need to perform studies that take into account location in the syllable or other higher-order prosodic domains such as the word, metrically strong positions in the prosodic hierarchy, or initial and final positions of prosodic domains other than the syllable. Additional perceptual assimilation data are needed to shed light on what constitutes the initial state of the L1 English–L2 German phonological IL, including how it is complicated by factors such as position-sensitivity, variable perceptual assimilation of the same phone to different L1 categories within the same position, and, potentially, by individual differences between learners' perceptual assimilation tendencies. In the end, there is likely a set of probable initial states, rather than one initial state, that will be identified as the starting point(s) for L2 German phonological acquisition as it relates to these fricatives.

Data Accessibility Statement

Additional files for this article, including stimuli, experiment code, and anonymized data set, may be accessed through the Open Science Framework here: <https://osf.io/gnyxu/>

Additional files

The additional files for this article can be found as follows:

- **Appendix 1.** Language Background Questionnaire. DOI: <https://doi.org/10.16995/labphon.6428.s1>
- **Appendix 2.** Sample images of perceptual assimilation category selection screens. DOI: <https://doi.org/10.16995/labphon.6428.s2>

Ethics and consent

All procedures reported in this study were approved by the Institutional Review Board of Indiana University, Bloomington (Protocol #1403865786). Informed consent was obtained from all participants in this study.

Acknowledgements

The authors would like to thank Jeffrey Holliday for his consultation regarding task design in OpenSesame as well as Christiane Kaden and Franziska Krüger for their patient hours in a recording booth.

Competing interests

The authors have no competing interests to declare.

Author contributions

Both authors contributed to conceptualization of the experimental design and writing of this manuscript. The first author completed all data collection and analysis. This article is based on part of the first author's 2019 dissertation, supervised by the second author.

References

- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in*

- second language speech learning: In honor of James Emil Flege* (pp. 13–34). Philadelphia: Benjamins. DOI: <https://doi.org/10.1075/llt.17.07bes>
- Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.66). [Computer program]. Retrieved from <http://www.praat.org/>
- Broselow, E., & Finer, D. (1991). Parameter setting in second language acquisition. *Second Language Research*, 7, 35–59. Retrieved from <https://www.jstor.org/stable/43104418>. DOI: <https://doi.org/10.1177/026765839100700102>
- Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14, 136–193. Retrieved from <https://www.jstor.org/stable/43104581>. DOI: <https://doi.org/10.1191/026765898669508401>
- Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory* (pp. 4–63). Malden, MA: Blackwell.
- Cardoso, W. (2011). The development of coda perception in second language phonology: A variationist perspective. *Second Language Research*, 27, 433–465. Retrieved from <https://www.jstor.org/stable/43103874>. DOI: <https://doi.org/10.1177/0267658311413540>
- Cheng, B., & Zhang, Y. (2015). Syllable structure universals and native language interference in second language perception and production: Positional asymmetry and perceptual links to accentedness. *Frontiers in Psychology*, 6, 1801. DOI: <https://doi.org/10.3389/fpsyg.2015.01801>
- Cho, M.-H., & Chung, J. (2010). Cross-language perception by prosodic position. *Eoneohag: Journal of the Linguistic Society of Korea*, 57, 83–108.
- Cho, M.-H., & Lee, S. S. (2007). Category matching between English and Korean consonants in different prosodic environments. *영어영문학 [English Language and Literature]*, 53, 731–753. DOI: <https://doi.org/10.15794/jell.2007.53.5.003>
- Christie, J., & Klein, R. (1996). Assessing the evidence for novel popout. *Journal of Experimental Psychology: General*, 125, 201–207. DOI: <https://doi.org/10.1037/0096-3445.125.2.201>
- Daidone, D., Kruger, F., & Lidster, R. (2015). Perceptual assimilation and free classification of German vowels by American English listeners. In The Scottish Consortium for ICPHS 2015 (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: Glasgow University.
- Daidone, D., Kruger, F., & Lidster, R. (2019, August). *Non-native discrimination of vowels, consonants, and phonemic length is better predicted by perceptual similarity than by Perceptual Assimilation category types*. Paper presented at the meeting of New Sounds, Tokyo, Japan. Retrieved from <https://www.researchgate.net/publication/336473784>
- Davidson, L. (2011). Phonetic and phonological factors in the second language production of phonemes and phonotactics. *Language and Linguistics Compass*, 5, 126–139. DOI: <https://doi.org/10.1111/j.1749-818X.2010.00266.x>
- Davis, S., & Cho, M.-H. (2003). The distribution of aspirated stops and /h/ in American English and Korean: an alignment approach with typological implications. *Linguistics*, 41, 607–652. DOI: <https://doi.org/10.1515/ling.2003.020>

- de Jonge, C. E. (1995). *Interlanguage phonology: Perception and production* (Unpublished doctoral dissertation). Indiana University, Bloomington.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39*, 379–397. DOI: <https://doi.org/10.2307/3588486>
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C. & Mehler, J. (1999). Epenthetic Vowels in Japanese: a Perceptual Illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1568–1578. DOI: <https://doi.org/10.1037/0096-1523.25.6.1568>
- Eckman, F., & Iverson, G. K. (2013). The role of native language phonology in the production of L2 contrasts. *Studies in Second Language Acquisition*, *35*, 67–92. DOI: <https://doi.org/10.1017/S027226311200068X>
- Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*, 345–360. DOI: <https://doi.org/10.1016/j.wocn.2007.11.002>
- Faris, M. M., Best, C. T., & Tyler, M. D. (2018). Discrimination of uncategorised non-native vowel contrasts is modulated by perceived overlap with native phonological categories. *Journal of Phonetics*, *70*, 1–19. DOI: <https://doi.org/10.1016/j.wocn.2018.05.003>
- Flege, J. E. (1995). Second language speech learning. Theory, Findings, and Problems. In W. Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge, UK: Cambridge University. DOI: <https://doi.org/10.1017/9781108886901.002>
- Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, *26*, 1–34. DOI: <https://doi.org/10.1017/S0272263104026117>
- Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɪ/ and /i/. *Journal of the Acoustical Society of America*, *99*, 1161–1173. DOI: <https://doi.org/10.1121/1.414884>
- Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, *107*, 2711–2724. DOI: <https://doi.org/10.1121/1.428657>
- Hall, T. A. (2013). Alveolopalatalization in Central German as markedness reduction. *Transactions of the Philological Society*, *112*, 143–166. DOI: <https://doi.org/10.1111/1467-968X.12002>
- Hallé, P. A., Best, C. T., & Levitt, A. (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics*, *27*, 281–306. DOI: <https://doi.org/10.1006/jpho.1999.0097>
- Hallé, P. A., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 592–608. DOI: <https://doi.org/10.1037/0096-1523.24.2.592>
- House, D. (1996). Differential perception of tonal contours through the syllable. In *Proceedings of the Fourth International Conference on Spoken Language Processing, ICSLP '96* (pp. 2048–

2051). Retrieved from <http://www.isca-speech.org/archive>. DOI: <https://doi.org/10.1109/ICSLP.1996.607203>

Jackson, C. N., & O'Brien, M. G. (2011). The interaction between prosody and meaning in second language speech production. *Die Unterrichtspraxis/Teaching German*, 44, 1–11. DOI: <https://doi.org/10.1111/j.1756-1221.2011.00087.x>

Jannedy, S., & Weirich, M. (2017). Spectral moments vs discrete cosine transformation coefficients: Evaluation of acoustic measures distinguishing two merging German fricatives. *Journal of the Acoustical Society of America*, 142, 395–405. DOI: <https://doi.org/10.1121/1.4991347>

Jannedy, S., Weirich, M., & Helmeke, L. (2015). Acoustic analyses of differences in [ç] and [j] productions in Hood German. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, 328. Retrieved from <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0328.pdf>

Johnston, W. A., & Swarting, I. S. (1996). Reassessing the evidence for novel popout. *Journal of Experimental Psychology: General*, 125, 208–212. DOI: <https://doi.org/10.1037/0096-3445.125.2.208>

Johnston, W. A., & Swarting, I. S. (1997). Novel popout: an enigma for conventional theories of attention. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 622–631. DOI: <https://doi.org/10.1037/0096-1523.23.3.622>

Kabak, B., & Idsardi, W. J. (2007). Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech*, 50, 23–52. DOI: <https://doi.org/10.1177/00238309070500010201>

Kilpatrick, A. J., Bundgaard-Nielsen, R. L., & Baker, B. J. (2019). Japanese Co-Occurrence Restrictions Influence Second Language Perception. *Applied Psycholinguistics*, 40, 585–611. DOI: <https://doi.org/10.1017/S0142716418000711>

Kilpatrick, A. J., Kawahara, S., Bundgaard-Nielsen, R. L., Baker, B. J., & Fletcher, J. (2018). Japanese coda [m] elicits both perceptual assimilation and epenthesis. *Proceedings of the International Symposium on Applied Linguistics (ISAPh 2018)* (pp. 79–83). Retrieved from https://www.isca-speech.org/archive/isaph_2018/index.html. DOI: <https://doi.org/10.21437/ISAPh.2018-14>

Kleber, F., Lowery, M., & Stegmaier, R. (2018). The production and perception of the German /s, ç, ʃ/ contrast. In *Proceedings of the Conference on Phonetics & Phonology in German-speaking countries (P&P 13)*, Berlin, Germany. DOI: <https://doi.org/10.18452/18805>

Lee, S., & Cho, M.-H. (2006). A positional effect in the perception of English anterior obstruents. *Korea Journal of English Language and Linguistics*, 6, 849–867.

Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *Journal of the Acoustical Society of America*, 125, 2670–2682. DOI: <https://doi.org/10.1121/1.3224715>

Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, 36, 141–157. DOI: <https://doi.org/10.1016/j.wocn.2007.03.001>

- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, *94*, 1242–1255. DOI: <https://doi.org/10.1121/1.408177>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences (Version 2.8.0). *Behavior Research Methods*, *44*, 314–324. [Computer program]. Retrieved from <http://osdoc.cogsci.nl/>. DOI: <https://doi.org/10.3758/s13428-011-0168-7>
- Otake, T., Yoneyama, K., Cutler, A., & Van der Lugt, A. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, *100*, 3831–3842. DOI: <https://doi.org/10.1121/1.417239>
- Park, H., & de Jong, K. J. (2008). Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics*, *36*, 704–723. DOI: <https://doi.org/10.1016/j.wocn.2008.06.002>
- Park, H., & de Jong, K. J. (2017). Perceptual category mapping between English and Korean obstruents in non-CV positions: Prosodic location effects in second language identification skills. *Journal of Phonetics*, *62*, 12–33. DOI: <https://doi.org/10.1016/j.wocn.2017.01.005>
- Polivanov, E. (1931). La perception des sons d'une langue étrangère [Perception of sounds in a foreign language]. *Travaux du Cercle Linguistique de Prague*, *4*, 79–96.
- R: A Language and Environment for Statistical Computing [Computer language and software]. Retrieved from <https://www.R-project.org/>
- Rose, M. (2010). Differences in discriminating L2 consonants: A comparison of Spanish taps and trills. In Y. Watanabe, M. Prior, & S.-K. Lee (Eds.), *Selected proceedings of the 2008 Second Language Research Forum* (pp. 181–196). Somerville, MA: Cascadilla Proceedings Project.
- Rose, M., & Darcy, I. (2011). The effect of syllable stress and L2 experience on the cross-language identification of L2 Spanish consonants in English [Unpublished manuscript]. Department of Second Language Studies, Indiana University.
- Scott, J. H. G. (2019). *Phonemic and Phonotactic Inference in Early Interlanguage: Americans Learning German Fricatives in L2 Acquisition* [Unpublished doctoral dissertation]. Indiana University, Bloomington. Retrieved from <https://dissexpress.proquest.com/search.html> (Publication No. 13878342)
- Segui, J., Frauenfelder, U., & Hallé, P. (2001). Phonotactic constraints shape speech perception: Implications for sublexical and lexical processing. In E. Dupoux (Ed.), *Language, Brain, and Cognitive Development: Essays in honor of Jacques Mehler* (pp. 195–208). Cambridge, MA: MIT Press.
- Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff, & R. T. Oehrle, with F. Kelley, & B. K. Stephens (Eds.), *Language Sound Structure: Studies in Phonology Presented to Morris Halle by His Teacher and Students* (pp. 107–136). Cambridge, MA: MIT Press.

- Shea, C. E., & Curtin, S. (2010). Discovering the relationship between context and allophones in a second language: Evidence for distribution-based learning. *Studies in Second Language Acquisition*, 32, 581–606. DOI: <https://doi.org/10.1017/S0272263110000276>
- Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech productions can precede perception. *Applied Psycholinguistics*, 3, 243–261. DOI: <https://doi.org/10.1017/S0142716400001417>
- Smith, J., & Kochetov, A. (2009). Categorization of non-native liquid contrasts by Cantonese, Japanese, Korean, and Mandarin listeners. *Toronto Working Papers in Linguistics*, 34, 1–15.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39, 456–466. DOI: <https://doi.org/10.1016/j.wocn.2010.09.001>
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America*, 109, 1691–704. DOI: <https://doi.org/10.1121/1.1353594>
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26, 311–344. DOI: <https://doi.org/10.1006/jpho.1998.0078>
- Strange, W., Bohn, O.-S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America*, 115, 1791–1807. DOI: <https://doi.org/10.1121/1.1687832>
- Streeter, L. A., & Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *Journal of the Acoustical Society of America*, 65, 1533–1541. DOI: <https://doi.org/10.1121/1.382917>
- Stevens, P. (1960). Spectra of fricative noise in human speech. *Language & Speech*, 3, 32–49. DOI: <https://doi.org/10.1177/002383096000300105>
- Trapman, M., & Kager, R. (2009). The acquisition of subset and superset phonotactic knowledge in a second language. *Language Acquisition*, 16, 178–221. DOI: <https://doi.org/10.1080/10489220903011636>
- Tronnier, M., & Dantsuji, M. (1993). An acoustic approach to fricatives in Japanese and German. In *Proceedings of the 3rd European Conference on Speech Communication and Technology (Eurospeech 1993)*, 271–274. Retrieved from https://www.isca-speech.org/archive/eurospeech_1993. DOI: <https://doi.org/10.21437/Eurospeech.1993-83>
- Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics*, 35, 49–64. DOI: <https://doi.org/10.3758/BF03205924>
- Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, 50, 351–360. DOI: <https://doi.org/10.3758/BF03212227>
- Wiese, H. (2006). „Ich mach dich Messer“: Grammatische Produktivität in Kiez-Sprache („Kanak-Sprak“) [„I make you knife“: Grammatical productivity in hood-language („Kanake-language“)]. *Linguistische Berichte [Linguistic Reports]*, 207, 245–273. Retrieved from https://www.uni-potsdam.de/fileadmin/projects/dspdg/Publikationen/Wiese2006_Messer.pdf

Winters, S. (2001). VCCV Perception: Putting place in its place. In E. Hume, N. Smith, & J. van de Weijer (Eds.), *HIL Occasional Papers: Surface Syllable Structure and Segment Sequencing, 4* (pp. 230–247). Leiden, Netherlands: Holland Institute of Generative Linguistics. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.532.9665>

Yoshida, K., & Hirasaka, F. (1983). The Lexicon in Speech Perception. *Sophia Linguistica: Working Papers in Linguistics Tokyo, 11*, 105–116.

