Original Article

The Relationship between Non-Native Perception and Phonological Patterning of Implosive Consonants

Madeleine Oakley [iD]
North Carolina State University, USA

Hannah Sande [iD]
University of California, Berkeley, USA

Abstract

This study uses non-native perception data to examine the relationship between perceived phonetic similarity of segments and their phonological patterning. Segments that are phonetically similar to one another are anticipated to pattern together phonologically, and segments that share articulatory or acoustic properties are also expected to be perceived as similar. What is not yet clear is whether segments that pattern together phonologically are perceived as similar. This study addresses this question by examining how L1 English listeners and L1 Guébie listeners perceive non-native implosive consonants compared with plosives and sonorants. English does not have contrastive implosives, whereas Guébie has a bilabial implosive. The bilabial implosive phonologically patterns with sonorants in Guébie, to the exclusion of obstruents. Two perception experiments show English listeners make more perceptual categorization errors between implosives and voiced plosives than Guébie listeners do, but both listener groups are more likely to classify implosives as similar to voiced plosives than sonorants. The results also show that Guébie listeners are better at categorizing non-native implosive consonants (i.e., alveolar implosives) than English listeners, showing that listeners are able to extend features or gestures from their L1 to non-native implosive consonants. The results of these experiments suggest a cross-linguistic perceptual similarity hierarchy of implosives compared with other segments that are not affected by L1 phonological patterning.

Corresponding author:
Madeleine Oakley, North Carolina State University, Department of English, Tompkins Hall, 2211 Hillsborough St, Raleigh, NC, 27607, USA.
Email: mo643@georgetown.edu

Introduction

The purpose of this paper is to investigate non-native speakers' perceptual categorization of implosive consonants to better understand the relationship between perceptual similarity and phonological patterning. Sounds that are perceived as similar (or are perceptually assimilated to one L1 category) often share phonetic properties. Phonetic similarity, defined by shared acoustic and/or articulatory features, also plays a key role in determining whether sounds pattern together phonologically (as classes of sounds that trigger or are triggered by the same phonological processes). What is unknown is whether phonological patterning influences the perceived similarity of segments. If two sounds pattern together in a language, are these sounds likely to be perceived as similar? Or is phonetic similarity alone responsible for perceived similarity of segments?

The studies presented here are designed to address these questions by examining how speakers of Guébie (a Kru language spoken in Southwest Côte d'Ivoire) differ from speakers of English in their perception of implosive consonants. Previous research on the perception of implosive consonants has shown that L1 English listeners have a tendency to perceive /ɓ/ as similar to /b/, showing that naïve listeners perceive implosives as similar to plosives (C. T. Best et al., 2001). Bilabial implosives and plosives are produced with a labial closure, and thus are produced with the same articulators. Phonetic similarity, defined by shared articulatory gestures and constriction location, is argued to play a key role in why L1 English listeners perceive non-native /ɓ/ as similar to /b/.

Phonetically similar segments also typically pattern together phonologically (Hansson, 2001; Mielke, 2005, 2012; Rose & Walker, 2004). Based on this observation, one might hypothesize that implosives will pattern with obstruents across languages. In fact, most feature theories assume that implosives share phonological features with obstruents (Greenberg, 1970; Hall, 2007; Halle & Stevens, 1971; Keating, 1984; Lombardi, 1995). However, implosives pattern with obstruents (rather than sonorants) in only a third of the world's languages that have implosives (Sande & Oakley, 2021). Whether patterning together phonologically affects the perceived similarity of two segments remains unknown. The present study examines whether the presence of contrastive implosives that pattern with sonorants in a speaker's L1 affects the perceived similarity of non-native implosives as obstruents versus sonorants. These findings have implications for the universal perception of implosive consonants and the representation of sounds.

1.1 Non-native perception

It has long been established that a speaker's L1 phonological system impacts their perception of sound contrast (C. T. Best et al., 2001; Flege, 1995; Hume & Johnson, 2001; Strange, 2011). If a listener's L1 contrasts two particular sounds, they can perceive the difference between the two segments. If a listener's L1 lacks the relevant contrast *and* the two phones are phonetically similar to a single L1 phone, then listeners may have difficulty perceiving the contrast between the two phones. As such, models of non-native speech emphasize how a speaker's L1 phoneme inventory influences the perception of non-native phones (C. T. Best, 1995; C. T. Best & Tyler, 2007; Flege, 1995; Flege & Bohn, 2021).

In this section, we focus on how the perceptual assimilation model (PAM) predicts non-native listeners will perceive the contrast between implosives and plosives. We set aside other models of non-native speech perception which focus on L2 speech perception (Flege, 1995; Van Leussen & Escudero, 2015) or infant speech perception (Kuhl et al., 2008) rather than adult naïve perception. According to PAM, a listener perceives non-native phones in relation to L1 phoneme categories, and a listener's ability to discriminate non-native phones depends on how these sounds are mapped

to L1 phones (C. T. Best, 1995). Different types of assimilation patterns predict how a listener discriminates two non-native phones.

There are three ways a non-native phone may be perceived in relation to L1 phones: a phone may be categorized as an L1 phone, it can be perceived as different enough from any L1 phone such that it is considered "uncategorizable," or it may be perceived as a non-speech sound (C. T. Best, 1995; C. T. Best et al., 2001; Faris et al., 2016). PAM makes explicit predictions about non-native perception of *contrast*, and therefore, the predictions for how a non-native contrast is perceived depend on the perceptual mapping between the non-native contrast and L1 categories. First, if two non-native phones are mapped to two separate phoneme categories in their L1, referred to as two-category (TC) assimilation, a listener will be excellent at perceiving the contrast between the two phones. Conversely, if two sounds are mapped equally well or poorly to the same category (SC assimilation), a listener will display poor discrimination abilities of the contrast. Listeners may also be sensitive to within-category phonetic differences of a non-native contrast. In this case, non-native phones are mapped to the same L1 category, but are not mapped equally well. In this situation, referred to as category goodness (CG) assimilation, listeners are predicted to show better discrimination of the non-native contrast than in SC assimilation, but show worse discrimination of the contrast than in TC assimilation. Finally, listeners are predicted to be excellent at discriminating the contrast between a non-native phone that is uncategorized compared with a phone that is categorized (UC assimilation) (Faris et al., 2016; Tyler et al., 2014).

Important for the present discussion, the non-native perception of implosive consonants has been analyzed in this framework. In a study examining how L1 English listeners perceive the implosive consonant, Best et al. (2001) find that listeners have difficulty perceiving the contrast between /ɓ/ and /b/. Listeners who have an implosive-plosive contrast in their L1 are not examined in this study, but it is assumed they would discriminate this contrast consistently. Listeners who had difficulty perceiving the contrast between /ɓ/ and /b/ also used similar descriptors for /ɓ/ and /b/ with consistent ratings, providing evidence that English listeners assimilated /ɓ/ and /b/ to their L1 English /b/ category (SC assimilation). Best et al. (2001) argue that the listeners who did not perceive the contrast between the bilabial implosive and the bilabial plosive only perceived the shared articulatory gestures between /ɓ/ and /b/, and failed to detect the gestures that are phonologically irrelevant in English (in this case, the lowering of the larynx during /ɓ/). A subgroup of listeners in the study were better at perceiving the /ɓ/-/b/ contrast than the SC assimilation group. This group of listeners also provided written descriptors that indicated they perceived phonetic differences between /ɓ/ and /b/ (CG assimilation). It is argued that this CG assimilation group did perceive the articulatory differences between /ɓ/ and /b/, unlike the SC assimilation group.

Perceived phonetic similarity between phones is a key component of predicting how non-native listeners will discriminate those phones (C. T. Best, 1995; C. T. Best & Tyler, 2007; Flege, 1995; Flege & Bohn, 2021; Strange, 2007, 2011). However, as discussed by Strange (2007), it is important to define phonetic similarity independently from perceptual similarity for models to have predictive power. PAM defines perceptual similarity in terms of shared articulatory organs between the phones, and the gestures of those organs. Similarity can also occur for a single gesture produced by different organs, such as a similar critical narrowing gesture produced across fricatives at different places of articulation (i.e., using different organs).

PAM adopts a direct realist perspective of speech perception, meaning the object of speech perception is defined as the articulatory gesture. Therefore, predictions for how listeners discriminate contrast between non-native phones depend on the shared articulatory organs and the gestures of the phones. Although there is variation across languages for how implosive consonants are produced (Ladefoged, 1968), /ɓ/ and /b/ are typically both produced with a complete closure at the labial position, and differ in the position of the larynx; the larynx is lowered in the production of

/ɓ/, which decreases air pressure. However, listeners who perceive only the shared articulatory organs and articulatory gestures of the bilabial implosive and plosive will perceptually assimilate /ɓ/ and /b/ to the same L1 category, /b/. Although the purpose of this paper is not to compare whether shared articulatory gestures or acoustic properties better predict perceived similarity, we will see that gestural accounts can explain the patterns of perception observed in this study.

Beyond the segmental level, there is also evidence that suggests that non-native listeners generalize contrastive cues in their L1 to discriminate non-native phones. For example, Pajak and Levy (2014) found that speakers who have a length contrast for vowels, but not consonants, in their L1 will be better able to perceive the contrast between long and short consonants than speakers who do not have the contrastive length for either consonants or vowels. Although listeners do not have a length contrast in their L1, having a length contrast for vowels facilitates listener perception of consonant length. Similarly, Bohn and Best (2012) found that contrastive rounding for L1 vowels facilitates listeners' perception of the non-native glide contrast between /w/ and /j/. Having rounding as a phonetic correlate to an L1 contrast aids listeners in their perception of non-native contrasts that also use the rounding cue. Taken together, these findings show that listeners are better at perceiving non-native contrasts if they have a native contrast that exploits the same phonetic correlates. For the current study, we might thus expect listeners to be better at perceiving the contrast between implosives and plosives at multiple places of articulation (such as alveolar and bilabial) even if they only have the implosive/plosive contrast at one place of articulation (bilabial).

This section has discussed how phonetic similarity between implosives and plosives interacts with the perception of these segments and how models of non-native perception predict the patterns observed. The experiments included in the present study are designed to further explore the perceived similarity of non-native plosives and implosives. For speakers whose L1 has the plosive-implosive contrast, do they consistently discriminate non-native plosives from implosives? Do they perceive non-native implosives as similar to plosives, despite phonological alternations in their L1 that suggest implosives pattern as sonorants? Section 1.2 discusses the phonological patterning of implosives, and how we expect phonetic similarity to play a role in phonological patterning.

1.2 Phonological patterning of implosives

Phonetic similarity is argued to play a key role in whether sounds pattern together phonologically. For example, sounds that have phonetically motivated features in common are more likely to be subject to vowel or consonant harmony (Rose & Walker, 2004, p. 476), or to dissimilation patterns (Bennett, 2013), as formalized in the agreement by correspondence model. Similarly, rules and constraints that refer to phonetically motivated distinctive features predict that phonetically similar sounds will pattern together. For example, in Modern Greek, velar sounds, which share a [PLACE] feature, pattern together phonologically in undergoing palatalization before non-back vowels. The class of non-back vowels (which are described as sharing a [-BACK] feature) patterns together in triggering this process.

In most feature theories, individual distinctive features are phonetically defined, grounded in acoustics, articulation, or perception (Chomsky & Halle, 1968; Clements & Hume, 1995; Jakobson et al., 1951). Even in emergent feature theory, phonetic similarity is predicted to motivate phonological natural classes (Mielke, 2004, 2008). Thus, we expect implosives to phonologically pattern together with sounds that are phonetically similar.

In defining phonetic similarity, Mielke (2012) uses acoustic and articulatory data to show how obstruents and sonorants differ in their phonetic properties. The results of this study indicate that

obstruents share articulatory and acoustic properties with one another, and sonorants share articulatory and acoustic properties with one another. Furthermore, segments which sometimes pattern with obstruents but sometimes pattern with sonorants, such as fricatives, are more phonetically similar to sonorants than other obstruents are (Mielke, 2012). These results importantly suggest that phonetic similarity is associated with which sounds pattern together phonologically and show that obstruent and sonorant classes are typically phonetically grounded.

Implosives are often assumed to be obstruents in the feature theory literature, perhaps because of certain phonetic similarities to plosives. For example, Greenberg (1970) states, "the phonological opposition in individual languages between ejectives and injectives (implosives) applies effectively only to obstruents, and is neutralized for sonorants and semi-vowels," and "the typical injective (implosive) obstruent is, on the other hand, a voiced stop" (pp. 123–124). Binary features such as [+CONSTRICTED] or privative features such as [CONSTRICTED GLOTTIS] have been used to distinguish implosives from obstruents and sonorants (Halle & Stevens, 1971; Keating, 1984; Lombardi, 1995). These approaches imply that implosives share all the features of stops, with an additional laryngeal feature.

Despite the assumption that implosives are obstruents, there has been very little work on the phonological patterning of implosives across languages. However, there are a number of descriptions of implosive patterning in individual languages, which show that implosives show similar phonological patterns to obstruents in some languages, but show similar phonological patterns to sonorants in others.

One language family in which implosives pattern phonologically with sonorants is Kru (Niger-Congo, Liberia, and Côte d'Ivoire) (Kaye, 1981; Marchese, 1979; Sande, 2017). For example, in Guébie (Eastern Kru, Côte d'Ivoire), there is one implosive sound, /ɓ/, which contrasts with bilabial stops /p/ and /b/ and sonorants /j/, /w/, and /l/. The implosive shows the same phonological distribution and alternation behavior as sonorants, while these same phonotactic restrictions and alternations do not apply to obstruents (Sande, 2017). Specifically, non-nasal sonorants and the implosive do not follow nasal consonants within a word, while obstruents regularly do (1). In addition, implosives and glides can be inserted to break up vowel hiatus (2). Note that while it is possible to produce the form in (2a) with either a [w] or [ɓ] between the underlyingly adjacent vowels, /w/ and /ɓ/ are otherwise contrastive in the language: /wi ^{3}/ "cry," /ɓi $^{3.1}$/ "plates." And most words of the shape CVLV or CVɓV can optionally surface as CCV, while other CVCV words cannot (3). In the Guébie examples below, tone is marked with superscript numerals where 4 is high and 1 is low.

1. Implosives and sonorants do not follow nasal consonants in Guébie

	C2 = Nasal	C2 = Sonorant	C2 = Implosive	C2 = Obstruent
C1 = Non-nasal	kaɲɪ$^{3.1}$ "mosquito"	kalɛ$^{2.3}$ "anger"	daɓara$^{4.4.4}$ "market"	dakɔ$^{3.1}$ "find"
C1 = Nasal	ŋʷɔnɔ$^{4.4}$ 'woman'	–	–	ɲɛpɛ$^{4.2}$ "sweep"

2. Guébie vowel hiatus avoidance

	Underlying VV	Surface VCV	Translation
a.	/dibo-ə$^{3.1.2}$/	[dibowə$^{3.1.1}$] or [diboɓə$^{3.1.1}$]	"the plantain"
b.	/ɟɔkʷɪ-a$^{2.3.3}$/	[ɟɔkʷɪja$^{2.3.3}$]	"the bird"

3.   Guébie CVCV ∼ CCV alternations

	CVCV	CCV	Translation
a.	jɪla [2.3]	jla [23]	"ask"
b.	bala [3.3]	bla [3]	"hit"
c.	duɓu [3.3]	dɓu [3]	"mourn"
d.	bete [3.1]	*bte [31]	"break"

The behavior of implosives in Guébie contrasts with languages like Hausa (Newman, 2000) or Fula (Paradis, 1992), where implosives pattern exclusively with obstruents and never with sonorants. For example, in Hausa, obstruent clusters surface as geminates with the features of the second obstruent (4a). Clusters of implosives and obstruents show the same behavior as obstruent-obstruent clusters (4b); however, clusters containing a sonorant and an obstruent do not alternate to surface as a geminate (4c). Implosives and obstruents both undergo the same assimilation process to the following consonant (4a, b), while sonorants do not (4c).

4.   Hausa cluster simplification

	Underlying	*Surface*	*Gloss*
a.	zàaf-zàafáa	zàazzàafáa	"hot"
b.	káɗ-kàɗáa	kákkàɗáa	"keep beating"
c.	fàrkáa	fárkáa, *fákkáa	"paramour"

In addition, in Hausa, coronal sonorants do not co-occur within a word; however, coronal implosives and obstruents freely co-occur with other coronals (cf. ɗàri, "hundred"). This is another example of how implosives pattern unlike sonorants and like obstruents in Hausa phonology.

In Kru languages like Guébie, implosives seem to pattern exclusively with sonorants, perhaps unexpectedly given previously proposed features of implosives. Hausa is an example of the expected case, where implosives pattern as a natural class with obstruents, to the exclusion of sonorants.

In a typological survey of 95 languages, Sande and Oakley (2021) show that among languages with contrastive implosive consonants as well as voiced obstruents and sonorants, implosives pattern phonologically with obstruents in 40%, and with sonorants in 38%. Languages can also show mixed behavior, in which implosives share certain phonological behaviors with both obstruents and sonorants (cf. Ikwere [Clements & Osu, 2002]). The latter type of language accounts for 21% of languages with contrastive implosives in the sample. If implosives only share phonological features with obstruents and not sonorants, it would be surprising to find such an even distribution of languages where implosives form a phonological class with sonorants as opposed to obstruents.

Section 1.1 summarized work which shows that phonetic similarity often leads to perceptual similarity. This section summarized work which shows that phonetic similarity also correlates with shared phonological patterning, but also discussed the case of implosives, which show mixed phonological behavior. What is not clear is the relationship between shared phonological patterning and perception. Do speakers perceive sounds as similar if they pattern together phonologically? Or does phonetic similarity alone impact whether listeners perceive two sounds as similar?

One reason for the gap in the literature regarding the link between phonological patterning and perception may be because segments that pattern together phonologically often share phonetically grounded features. Implosives provide an opportunity to examine the relationship between phonological patterning and perception because they are phonetically similar to obstruents (namely, shared articulatory gestures and organs), but pattern phonologically with sonorants in many languages (Sande & Oakley, 2021). The primary objective of the current experiments is thus to investigate whether listeners perceive two sounds as similar if they pattern together phonologically in their L1.

2 Hypotheses

The findings of Best et al. (2001) showed that non-native listeners perceive non-native implosives and plosives as similar, detecting the shared articulatory organs and gestures between the segments. Because implosives share primary articulatory organs and gestures with plosives, while most sonorants do not share the same articulatory gestures, we expect implosives to be perceived as more similar to plosives. However, if the perceived similarity is influenced (in part) by phonological patterning, we may expect listeners who have implosives that pattern with sonorants in their L1 to perceive implosives as similar to sonorants. As such, the present experiments examine how listeners perceive the non-native contrast between implosives and plosives/sonorants.

To investigate whether the presence of a sonorant-patterning implosive consonant impacts the non-native perception of implosives as similar to plosives/sonorants, two listener groups completed two perception experiments; the first listener group consists of L1 English speakers who have no exposure to implosive consonants, and the second listener group consists of L1 Guébie speakers, who have an implosive consonant in their L1 inventory. English listeners are included in this study as naïve listeners who have no exposure to the implosive-plosive contrast. Although it would be ideal to test a listener group who has an implosive-plosive contrast in their L1 *and* where implosives pattern phonologically with obstruents (as in Hausa, mentioned above), limitations to data collection during a global pandemic prohibited this possibility. However, English listeners will serve as a baseline comparison for Guébie listeners, because English listeners do not have sonorant-patterning implosives, or in fact any contrastive implosive in their L1. The Guébie listeners are included in this study to investigate whether having sonorant-patterning implosives in a listener's L1 impacts the perceived similarity between non-native implosives and plosives or sonorants.

Experiment 1 is designed to examine whether speakers who have contrastive implosives in their L1 differ from naïve listeners in their perceptual categorization of implosives and plosives/sonorants. The two listener groups are expected to differ in their categorization errors for the bilabial implosive. Based on previous findings examining the perceptual categorization of bilabial implosives (C. T. Best et al., 2001), L1 English listeners are predicted to have poor discrimination of /ɓ/ and /b/. This finding would suggest that L1 English listeners are perceptually mapping the bilabial implosive consonants to their L1 bilabial plosive (SC assimilation, according to PAM), which leads listeners to have difficulty discriminating between the two phones. Because Guébie has contrastive bilabial implosive consonants, L1 Guébie listeners are predicted to correctly categorize /ɓ/ compared with plosives and compared with sonorants. Correct categorization of the bilabial implosives compared with other segments suggests that the stimuli are assimilated to separate L1 categories (TC assimilation, according to PAM), meaning that Guébie listeners are mapping the implosive /ɓ/ in the stimuli to their L1 implosive category, and the plosives and sonorants to their respective L1 categories.

Guébie only has one implosive segment in the phonological inventory (the bilabial implosive), but this study will also examine how Guébie and English listeners categorize the alveolar implosive compared with obstruents and sonorants. Extending the results from the bilabial implosive, L1 English listeners are anticipated to make categorization errors between /ɗ/ and /d/. This result would further support the assumptions of PAM, based on the fact that the primary place of articulation would impact listener perceptual assimilation. For the Guébie listeners, there are two possibilities for how they will perceptually categorize /ɗ/ compared with other segments. First, because Guébie does not have a contrastive alveolar implosive, there is a possibility that Guébie listeners will perform similarly to L1 English listeners. If Guébie listeners map /ɗ/ to their L1 /d/, then it is predicted that listeners will make categorization errors between /d/ and /ɗ/ (SC assimilation, according to PAM). This result would show that the presence of an implosive category in speakers' L1 does not impact the categorization of phones at other places of articulation, and furthermore would suggest that there is a cross-linguistic tendency for implosives to be categorized as plosives at the same place of articulation. The second possibility is that Guébie listeners will correctly categorize the alveolar implosive compared with plosives and sonorants, despite not having an alveolar implosive category in their L1. If the Guébie listeners correctly categorize /ɗ/, but the L1 English listeners do not, then the presence of /ɓ/ in their L1 inventory can be said to have influenced the perceptual assimilation of /ɗ/. Listeners would have to generalize the implosive feature, or articulatory gesture, from the bilabial implosive to the alveolar implosive.

Experiment 2 examines whether listeners with contrastive implosive consonants in their L1 differ from naïve listeners as to whether they perceive implosives to be more similar to plosives or sonorants. First, because L1 English listeners have a tendency to perceptually assimilate /ɓ/ to their L1 /b/ category (Best et al., 2001), it is hypothesized that the L1 English listeners in this study will select /b/ as the most similar sound to /ɓ/, and will select /d/ as the most similar to /ɗ/. This result would provide further evidence that L1 English listeners' categorization errors from Experiment 1 are caused by the perceptual assimilation of implosives to plosive categories, which implies that implosives are perceptually similar to plosives.

For the Guébie listeners, there are two possibilities. The first possibility is that listeners will perceive implosives as more similar to plosives. This result would show that implosive consonants are perceived as more similar to plosives, even in languages where implosives pattern phonologically as sonorants. Based on the fact that implosives and plosives are produced with similar articulatory organs and gestures, this finding would support theories of non-native speech perception that claim that articulatory gestures are the relevant object of perception. The second possibility is that listeners will perceive implosives as more similar to sonorants. This finding would show that the phonological patterning of implosives is correlated to the perceived phonetic similarity of implosives and sonorants.

3 Experiment 1: ABX task

3.1 Methods

This section investigates the perceptual properties of implosives as compared with obstruents and sonorants. We run a set of perceptual experiments with English speakers, who do not have an implosive category, and with Guébie speakers, in whose L1 implosives pattern with sonorants.

Two different listener groups completed Experiment 1. The first listener group consisted of 20 native English speakers (3 males and 17 females) who had no previous exposure to any language with implosive consonants. All listeners reported normal hearing and were between the ages of 18 and 22 years (mean = 18.8 years). Group 1 listeners received course credit for their participation.

The second group of listeners are native speakers of Guébie, and therefore have an implosive category in their L1. All Guébie participants were also familiar with French, to varying degrees. Eleven Guébie listeners (9 males and 2 females) completed Experiment 1 in Gnagbodougnoa, Côte d'Ivoire. Although the exact age of the participants is unknown due to cultural restrictions on asking participants to disclose their ages, all listeners were of working age, roughly between 18 and 45 years. The majority of the participants were judged to be in their mid-20s based on the work they did in the village. The imbalance in the number of L1 English participants and L1 Guébie participants was caused by inherent differences in the data collection sites and language communities. There are a limited number of eligible participants in Gnagbodougnoa. Many speakers were uncomfortable using a laptop and did not want to participate. Many others could not take the time away from work (including child care and cooking) to participate in the study. Many women would not participate in the study due to cultural expectations of the community. Despite the limitations related to running experiments in a rural community, this study provides novel perception data from listeners of an endangered language. The impact of the limitations on the interpretation of the results will be discussed further in Section 6.

The first listener group completed the experiment in a sound-attenuated booth on Georgetown University's campus. The second group of listeners completed the experiment in residences and school buildings. The conditions in which the English and Guébie experiments were carried out were necessarily different and impact how results are interpreted. English-listener experiments were carried out in a lab setting at Georgetown University. A laboratory-like setting was not available in Gnagbodougnoa, Côte d'Ivoire, where the Guébie-listener experiments were run. Instead, Guébie-listener experiments were run in homes or a school building, where there were consistent background noises as well as regular interruptions.

In Experiment 1, listeners performed an ABX categorization task to assess how participants categorized implosives compared with plosives and sonorants. Participants were presented with three stimuli in each trial and asked to categorize the third sound in the sequence as either the same as the first sound or the second sound. Experiment 2 is a similarity task designed to assess the perceived similarity between the segments and will be described further in Section 4. It is important to note here that it was not possible to ask participants to provide written responses about how they perceived the stimuli, which would provide further insight into the perceptual categorization of the sounds (see Best et al., 2001) for a description of these methods). Guébie does not have a writing system, and many speakers of Guébie cannot read or write any language. It is therefore not possible to assess perceptual categorization based on orthographic responses.

The stimuli were nonce words of the form [aCa], with obstruents, sonorants, and implosives in the C position. The consonants in the C position included [ɓ, b, p, m, ɗ, d, t, n, l, j, w], for a total of 11 nonce word stimuli. One female and one male speaker were recorded producing each of the stimuli, and each speaker was a trained linguist. The female speaker is a heritage speaker of Vietnamese. The male speaker is a native speaker of French and a speaker and researcher of Laal. These particular speakers were selected because Vietnamese and Laal both contain implosive consonants and because both speakers are trained linguists who were comfortable producing the nonce words. The speakers could not be Guébie speakers because this task is intended to compare how *non-native* listeners perceive the contrast between implosives and plosives/sonorants. In addition, the Guébie community is small enough that listeners would recognize and potentially be distracted by recognizing the voice of the Guébie speaker. The phonological patterning of implosives in Laal and Vietnamese and the potential effects this could have on the interpretation of the perception results are discussed in Section 6. Both speakers were between the ages of 22 and 35 years at the time of recording. In the recording sessions, the speakers repeated the nonce words three times. The middle token of each repetition was selected for the perception task.

Each trial consisted of three nonce words. The third word, X, was always of the same type as one of the first two words, A or B: $[aɓa]_A$ $[ala]_B$ $[aɓa]_X$. The X token was produced by the male speaker, and the A and B tokens were produced by the female speaker.

The experiment started with a practice phase of 12 trials to familiarize participants with the task. After the practice phase, participants began the experiment. Each consonant pair was compared in a counterbalanced order. This means that for each pair of stimuli, there were four trials: ABB, ABA, BAA, and BAB. This totals to 220 trials. To keep the experiment of reasonable length, particularly for the Guébie listener community who does not have familiarity with this type of task, the experiment was shortened in the following way: participants completed either the ABB/ABA trials or the BAA/BAB trials, with the exception of the trials that included an implosive consonant. [ɓ] and [ɗ] were included in all four counterbalanced orders for each participant. This means that each participant heard the trials for the "distractor" pairs, for example [ala] compared with [apa] in only ABB/ABA pairs, or BAA/BAB pairs, but the implosive words, [aɓa] [aɗa], were presented in ABB, ABA, BAA, and BAB orders for every pair comparison. This allowed the experiment to be of reasonable length while still collecting relevant responses from each participant. Each participant thus completed 148 trials in the testing phase of Experiment 1. The experiment lasted approximately 20 min.

Participants wore headphones and the stimuli were presented auditorially via PsychoPy on a laptop computer. At the start of the experiment, participants were instructed that they would hear a sequence of three nonce words, and would need to decide whether the third "word" was more similar to the first or second word in the sequence. They were instructed to press the "1" key on the computer if they thought the third word was more similar to the first word and to press the "2" key if they thought the third word was more similar to the second word. The researcher presented both written and oral instructions to the participants. During each trial, as each "word" was presented, numbers 1 and 2 lit up on the screen to inform participants when they were hearing the first versus second word.

There are a number of differences in the environments where the two groups performed the perception tasks that affected the procedures. For English participants, experiment instructions were presented in written English, on a computer screen, prior to the start of the task. Many Guébie participants were illiterate and could not read instructions on the computer screen. For this reason, instructions for Guébie speakers were explained orally as well as provided in written form.

For the Guébie-listener experiments, instructions were provided in French, the contact language of the area. These instructions were presented both orally and in written French at the start of the PsychoPy experiment. A number of the Guébie-speaking participants, however, were much more comfortable speaking and listening to Guébie than French. In such situations, it was necessary to ask another Guébie speaker who understood the task to be present to explain the instructions in Guébie, in addition to the French instructions. Because Guébie does not have an orthography, it was not possible to provide written instructions to listeners in Guébie. Participants were better able to understand the task when explained by a Guébie speaker who could answer their questions as they explained the instructions.

Despite oral and written instructions in French and Guébie, a number of Guébie-speaking participants had trouble understanding the task. Data from five Guébie speakers had to be excluded from the results due to a failure to understand the instructions. For example, one participant chose Sound B (pressed Number 2 on the keyboard) for every trial, and another alternated between Sounds A and B consistently throughout the full experiment, no matter which sounds were being compared or in which order. Data from one English speaker was deleted because the computer crashed and would not restart part way through the study. No other data were excluded from the analysis.

Mixed effects models were run to see what factors impacted correct versus incorrect categorization, with "correct response" as the dependent variable. "Correct response" is coded as a categorical variable, with "yes" coded for correct categorization responses and "no" coded for incorrect categorization responses. All fixed effects that are predicted to impact the correct categorization of each word were included in the original model. Subsequently, the Drop1 function in *R* (R Core Team, 2021) was used to compare models with fewer factors. Fixed effects included distractor word (categorical variable for which word is compared with the correct word), X-position word (categorical variable for the word that was tested), the interaction of the X-position word and the distractor word, whether the trial compared across only alveolar consonants (coded as a binary variable, with "yes" meaning all words in the A, B, or X position contained alveolar sounds and "no" meaning not all words in the A, B, or X position contained alveolar sounds), whether the trial compared only bilabial consonants (coded as a categorical variable, with "yes" meaning all words in the A, B, or X position contained bilabial sounds and "no" meaning not all words in the A, B, or X position contained bilabial sounds), whether the trial compared across place of articulation (coded as a binary variable with "yes" meaning not all words in the A, B, or X position contained target sounds with the same place of articulation and "no" meaning that all words did contain target sounds at the same place of articulation), whether the trial included an implosive consonant in the A, B, or X position (a binary variable with "yes" meaning that there was at least one word that contained an implosive consonant in the trial and "no" meaning there was not a word that contained implosive consonants in the trial), and whether the trial included an implosive in the X position (a categorical variable with "yes" indicating the word in the X position contained an implosive consonant and "no" indicating the X position word did not contain an implosive consonant). Random effects included recency (a categorical variable with "yes" meaning that the correct response word was played as the second sound in the sequence and "no" meaning the correct response word was played as the first word in the sequence), condition type (a categorical variable with type A for ABB/ABA trials or type B for BAA/BAB trials), participant, and additional languages that participants spoke (all categorical variables). "Recency" was included as a random effect to account for the fact that listeners may be more likely to correctly categorize a word if it is the same type as the second word in the sequence. "Condition type" is included to account for the fact that each participant heard only half of the counterbalanced trials. R defaults to mapping categorical levels onto the intercept in alphabetical order. The signs of the B estimates in the results section reflect this default mapping. The models were run in R using the GLMER function, with the family type as binomial (R Core Team, 2021).

Analyses were run on the two listener groups separately and together. In the model with all the listeners included together, the following fixed effects were included in the logistic regression model: the interaction between language background and whether the trial included implosive consonants, and the three-way interaction between language background, the X-position word, and the distractor word. Two separate models were also run for the two listener groups because of the different experimental conditions. Although the two listener groups completed the same perception task, there were many factors in addition to native language that could not be controlled for, which may have caused the two listener groups to perform differently. Running the analyses separately as well as together shows important differences across groups and experimental conditions.

All listener responses, PsychoPy scripts, R scripts, and stimuli recordings are available on the California Language Archive, bundles 2014-15.123 and 2014-15.224: http://dx.doi.org/doi:10.7297/X29C6W8R (Bodji & Sande, 2014).

Table 1. Listener Group Categorization Error Rates for All Trials and Trials that Included an Implosive Consonant.

	English listeners	Guébie listeners
Total number of trials	2,960	1,628
Percentage of implosive trials	51% (1,520)	51% (836)
Categorization error rate (count)	6.9% (207)	34.2% (558)
Percentage of errors with implosive (count)	73.4% (152)	53.1% (290)
Percentage of errors with [ɓ] (count)	41.5% (86)	28.4% (159)
Percentage of errors with [ɗ] (count)	35.2% (73)	25.8% (144)

3.2 Experiment 1 results

The data were analyzed first as a proportion of errors to gain a descriptive picture of how categorization errors are affected by including an implosive consonant in a trial. Categorization errors are defined by instances where participants incorrectly identified the X token as either the A or B token.

Table 1 shows the number of categorization errors made by each listener group. The total number of trials refers to all the trials for each participant group. The total percentage of implosive trials is the percent of all trials that contained an implosive in the X or A/B position, meaning trials where participants were categorizing implosives compared with plosives/sonorants, or vice versa. The categorization error rate refers to the number of incorrect responses for each listener group. The percentage of errors with implosives are the number of errors that occurred during a trial with an implosive consonant.

Overall, English listeners had a low rate of categorization errors, at 6.9%. However, of the 207 errors made by English listeners, 152 (73.4%) were in trials that included implosive consonants. Because 51.35% of total trials included an implosive consonant, implosive trials are disproportionately present in categorization errors (73.4%). Broken down by place of articulation of the implosive consonant, 41.5% of errors were made on trials that contained [ɓ], and 35.2% of errors were made on trials that contained [ɗ]. Note that trials may contain both [ɓ] and [ɗ], hence why the categorization error rate for each implosive does not add up to the total error rate for all implosive consonants.

Guébie listeners overall had a higher rate of errors, at 34.2%. This higher categorization error rate is likely due to experimental conditions in the field and limitations of running perceptual experiments with listeners who are unfamiliar with these types of tasks. However, of the categorization errors made by Guébie listeners, only 53.1% included implosives. Because 51.35% of trials include an implosive consonant, the proportion of errors for Guébie listeners that contained an implosive is similar to the overall distribution of implosives in the experiment, showing that, unlike the English speakers, Guébie listeners did not make more errors when implosives were present than otherwise. The number of errors made on trials with the bilabial and alveolar implosives are very similar for Guébie listeners. In all, 28.4% of errors were made on trials that contained [ɓ], and 25.8% of errors were made on trials that contained [ɗ].

Next, a logistic mixed-effects model was run for all of the listeners combined to determine which factors significantly affected correct categorization. The first model included all the fixed effects and random effects outlined in Section 3.1, using the following R syntax:

Table 2. English and Guébie Listeners in ABX Condition: Fixed Effects in the Output of Logistic Mixed-Effects Model with Correct Categorization as the Dependent Variable.

	B Estimate	SE	Z	p
(Intercept)	2.912	0.314	9.258	<2e−16***
Distractor = aɓa	0.305	0.324	0.940	.347
Distractor = ada	0.034	0.245	0.139	.889
Distractor = aɗa	0.318	0.344	0.925	.354
Distractor = aja	0.129	0.232	0.558	.576
Distractor = ala	0.122	0.233	0.523	.600
Distractor = ama	0.581	0.221	2.627	.0086**
Distractor = ana	0.625	0.257	2.425	.0153*
Distractor = apa	0.288	0.214	1.348	.177
Distractor = ata	0.53	0.255	2.077	.037*
Distractor = awa	0.594	0.246	2.415	.015*
SoundX = aɓa	0.076	0.322	0.238	.811
SoundX = ada	0.295	0.254	1.161	.245
SoundX = aɗa	0.352	0.347	1.014	.31
SoundX = aja	0.293	0.238	1.231	.218
SoundX = ala	−0.38	0.223	−1.699	.089
SoundX = ama	0.569	0.223	2.552	.010*
SoundX = ana	0.616	0.261	2.356	.018*
SoundX = apa	0.410	0.219	1.869	.061
SoundX = ata	0.787	0.266	2.949	.003**
SoundX = awa	0.141	0.233	0.606	.544
AlveolarsOnly = True	−0.671	0.205	−3.27	.0010 **
BilabialsOnly = True	−0.748	0.199	−3.749	.0001***
Implosive = True	−0.905	0.330	−2.741	.006**
Language = Guébie	−2.637	0.268	−9.820	<2e−16***
Implosive = True*Language = Guébie	1.011	0.196	5.144	2.69e−07***
Marginal R^2				.277

SE: standard error.
*$p \leq$ .05. **$p \leq$ .01. ***$p \leq$ .001.

glmer.**all**<−glmer (correct ~ distractor + SoundX + alveolars + bilabials + bilabial.alveolar + implosive + implosiveX + language + SoundX*distractor + language*implosive + language*SoundX*distractor + (1|recent) + (1|Participant) + (1|Condition), **family** ="binomial," **data** = abx.**all**)

The Drop1 function in R was used to determine which factors could be eliminated and improve model fit with the following R syntax:

drop1(abx.8.abx.all.glmer, test = "none")

Dropping the three-way interaction between language background, X-position word, and distractor word lowers the Akaike information criterion (AIC) from 3,455.8 to 3,402.4, and an analysis of variance (ANOVA) model comparison shows that the model *without* the three-way interaction has a significantly better fit (χ^2 = 162.56, df = 108, p = .0005***). The Drop1 function in R was
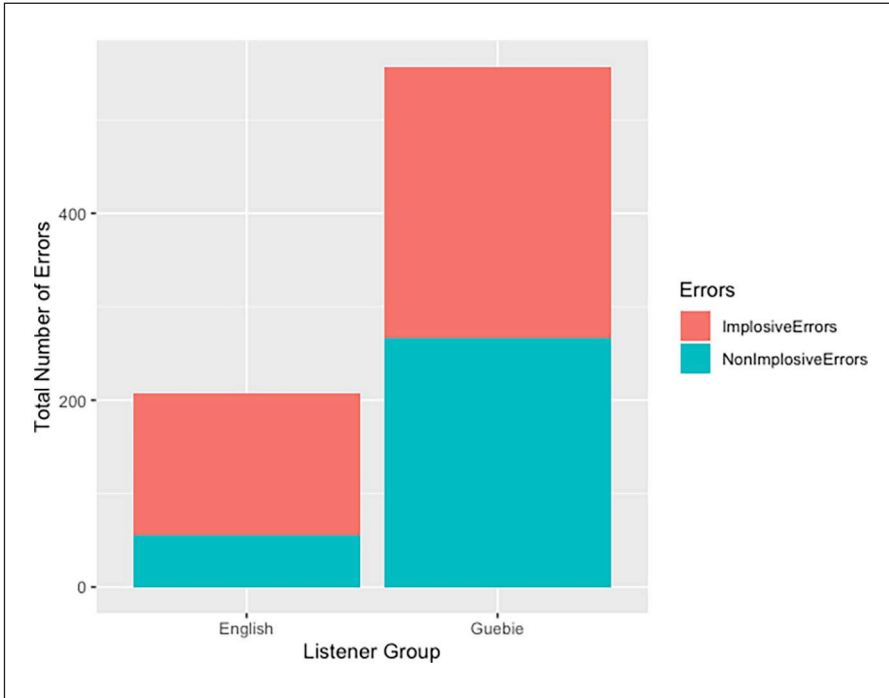
Figure 1. Total categorization errors for trials that contained implosive consonants and those that did not, by listener group.

used again to determine what other factors could be eliminated to improve the model fit of the new model. Eliminating the interaction between the X-position word and the distractor word lowers the AIC from 3,402.4 to 3,390.5, and again, an ANOVA model comparison shows that eliminating this interaction significantly improves the model fit ($\chi^2 = 160.11$, $df = 86$, $p = 2.18e{-}6***$). Finally, the Drop1 function in R was run on this model to determine whether any factors could be dropped to improve the model fit. There are no other fixed effects that can be dropped to lower the *AIC*. Table 2 shows the output of this final model. Nakagawa's marginal R^2 values were calculated using the tabmodel function (Lüdecke, 2021).

Results show that if the words [ama], [ana], or [ata] were in either the distractor position or in the X-word position, *correct* categorization responses were more likely. If [awa] was in the distractor position, *correct* categorization is significantly more likely. Alternatively, if trials only contained bilabial consonants or only contained alveolar consonants, categorization is significantly more likely to be *incorrect*. If trials contain implosives in any position, categorization is significantly more likely to be *incorrect*. Overall, Guébie listeners are more likely to make *incorrect* categorizations. This is unsurprising given the environment in which Guébie listeners performed the task.

Importantly, the interaction between an implosive being present in a trial and Guébie listeners is significant. Guébie listeners are more likely to make correct categorizations on trials that included implosive consonants than English listeners are. This interaction is seen in Figure 1, which shows the total errors made by each listener group, broken down by whether the trial contains an implosive consonant or not. As can be seen, Guébie listener errors are evenly distributed between trials

that contained implosive consonants and those that did not, whereas the English listener errors disproportionately occur in trials that contain implosive consonants.

A majority of the categorization errors made during implosive consonant trials by the English listener group occurred when implosives were co-present with voiced plosives at the same place of articulation. Of the 86 errors made in trials containing the word [aɓa] in either the A or B position, 38 errors were made when the comparison word was [aba]. The next highest number of errors was made when comparing [aɓa] to [apa], which only had a total of 10 errors for all English listeners. The alveolar implosive shows similar categorization errors. Of the 73 errors that English listeners made on trials containing the word [aɗa], 30 errors were made when the comparison word was [ada]. The next highest number of errors was made when comparing [aɗa] to [aba] or [ata], both of which only contained eight total errors for all English listeners. These results show that when English listeners made categorization errors with implosive consonants, they were usually miscategorizing implosives as voiced plosives at the same place of articulation.

Next, a logistic regression model was run for each listener group separately to determine what factors impacted correct categorization for each speaker group. First, for the English listener group, a model was run with all of the factors described in Section 3.1 included. The following R syntax was used to run this model:

abx.Eng.glmer<–glmer(correct distractor + SoundX + alveolars + bilabials + bilabial.alveolar + implosive + implosivex + SoundX*distractor + (1|recent) + (1|Participant) + (1|Condition), **family** = "binomial," **data** = abx.Eng)

The Drop1 function in R was used to determine what factors can be excluded to improve model fit. Excluding the interaction between the X-position word and the distractor word lowers the *AIC* from 1,373.1 to 1,365.5. Comparing the two models with the ANOVA function shows that dropping this interaction significantly improves model fit (χ^2 = 164.37, *df* = 86, *p* = 7.61e−7***). Dropping the factor for "implosive" lowers the *AIC* from 1,365.5 to 1,363.7. The model without the implosive factor is not a significantly better fit than with the implosive factor (χ^2 = .225, *df* = 1, *p* = .6353). Dropping any other factor does not improve model fit. The model with the following R syntax was therefore used for the English-listener responses:

abx.Eng.glmer<–glmer(correct distractor + SoundX + alveolars + bilabials + bilabial.alveolar + implosive + implosivex + (1|recent) + (1|Participant) + (1|Condition), **family** = "binomial," **data** = abx.Eng)

The output of this model for English listeners is seen in Table 3.

English listeners are more likely to correctly categorize sounds when the distractor word is [ama], [ana], or [apa]. English listeners are also significantly more likely to make correct categorizations when the X-position word is [ama], [apa], or [ata]. English listeners are significantly *less likely* to make correct categorizations when the X-position word is [ala]. Perhaps unsurprisingly, English listeners are also less likely to make correct categorizations when the trial contains only alveolar consonants or when the trial contains only bilabial consonants, meaning categorization errors are more likely within trials at the same place of articulation.

Next, separate logistic mixed-effects models were run for the Guébie listener group to determine which factors had a significant impact on correct classification for Guébie listeners. The first model was run using the following R syntax, and all the factors are described in Section 3.1:

Table 3. English Listeners in ABX Condition: Significant Predictor Variables in the Output of Logistic Mixed-Effects Model with Correct Categorization as the Dependent Variable.

	B Estimate	*SE*	*Z*	*p*
(Intercept)	2.999	0.464	6.462	1.03e−10***
Distractor = aɓa	1.301	0.495	0.607	.543
Distractor = ada	−0.159	0.401	−0.398	.690
Distractor = aɗa	0.013	0.547	0.025	.980
Distractor = aja	0.140	0.384	0.364	.715
Distractor = ala	0.300	0.411	0.731	.464
Distractor = ama	1.674	0.396	4.226	2.38e−05***
Distractor = ana	0.960	0.474	2.023	.043*
Distractor = apa	1.447	0.375	3.857	.00011***
Distractor = ata	0.831	0.456	1.820	.068
Distractor = awa	0.580	0.433	1.340	.180
SoundX = aɓa	−0.395	0.510	−0.776	.437
SoundX = ada	−0.442	0.444	−0.997	.318
SoundX = aɗa	−0.627	0.572	−1.096	.273
SoundX = aja	0.295	0.491	0.601	.547
SoundX = ala	−0.821	0.387	−2.118	.034*
SoundX = ama	1.413	0.442	3.197	.0013**
SoundX = ana	0.098	0.471	0.208	.835
SoundX = apa	1.047	0.402	2.60	.009314**
SoundX = ata	1.423	0.633	2.245	.0247*
SoundX = awa	−0.239	0.427	−0.561	.574
AlveolarsOnly = True	−0.943	0.362	−2.600	.009316**
BilabialsOnly = True	−2.149	0.367	−5.849	4.93e−09***
Implosive = True	−0.228	0.482	−0.474	.635
Marginal R^2				.253

SE: standard error.
*$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

abx.Gue.glmer<−glmer(correct distractor + SoundX + alveolars + bilabials + bilabial.alveolar + implosive + implosivex + SoundX*distractor + (1|recent) + (1|Participant) + (1|Condition), **family** = "binomial," **data** = abx.Gue)

Again, the Drop1 function in R was used to determine whether excluding any factors can improve model fit. Excluding the interaction between the X-position word and the distractor word reduces the model's *AIC* from 2,089.7 to 1,999.3. This does not significantly improve the model fit ($\chi^2 = 81.60$, $df = 86$, $p = .61$). Because of the length of the output of the first fixed-effects model, a summary of the significant factors at the .05 level predicting correct categorization for Guébie listeners is presented in Table 4. See the California Language Archive (Bodji & Sande, 2014) for the complete results.

These results show that Guébie listeners are more likely to correctly categorize glides and liquids when they are compared with alveolar stops. Specifically, [aja] is more likely to be correctly identified when compared with words containing the voiced and voiceless alveolar stops. [ala] is more likely to be correctly identified when compared with [ada]. This suggests that Guébie listeners are more likely to make correct categorization when approximants are compared with alveolar

Table 4. Guébie Listeners in ABX Condition: Significant Predictor Variables in the Output of Logistic Mixed-Effects Model with Correct Categorization as the Dependent Variable.

	B Estimate	SE	Z	p
Distractor = ada*Sound X = aja	3.986	1.805	2.208	.027*
Distractor = ata*Sound X = aja	3.721	1.607	2.314	.020*
Distractor = ala*Sound X = aja	3.724	1.850	2.013	.044*
Distractor = ada*Sound X = ala	3.986	1.805	2.208	.027*
Distractor = aɓa*Sound X = awa	−2.388	1.189	−2.007	.044*
Marginal R^2				.341

SE: standard error.
*$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

plosives. Listeners are also more likely to make correct categorizations when [aja] is compared with [ala]. Although the question of how alveolar approximants and plosives are perceived when compared with each other is not central to the discussion of this paper, Table 4 importantly shows that listeners are less likely to make categorization errors across places of articulation for approximants and plosives.

Interestingly, Table 4 shows that Guébie listeners are likely to make categorization errors when the labiovelar glide is compared with the bilabial implosive. [w] and [ɓ] are contrastive in Guébie, but listeners often make categorization errors when [awa] is in the X position of nonce words and is compared [aɓa]. Implosives were not otherwise significantly miscategorized by Guébie speakers.

4 Experiments 2 and 3: ABC and mixed conditions

4.1 Experiment 2 methods

Experiment 2 was designed to test more directly whether participants perceive implosives as more "similar" to plosives or sonorants. Rather than a categorization task, Experiment 2 was a forced choice similarity task, in which participants identified a nonce word as more "similar" to one word in a sequence than another. We refer to this forced choice similarity task as the ABC condition.

Another note must be made here as to why the ABC task was necessary to determine how "similar" the listeners found implosive consonants to be compared with sonorants and obstruents. Again, while other studies examining non-native perceptual assimilation of sounds have relied on tasks that use orthography to examine which L1 sounds the non-native sounds are mapped to (C. T. Best et al., 2001; Strange et al., 1998; Tyler et al., 2014), it is not possible to perform such tasks with Guébie listeners. An example of such a task would be to ask listeners to write down the sounds (or "words") they hear after the ABX task, and to give each transcription a rating for how "good" the fit between the stimuli and the orthography are. There is no orthography for Guébie, and many of the participants have low literacy. The ABC task is therefore developed to examine the perceived perceptual similarity of segments without orthography and provides an avenue to test the non-native perception on a wider range of languages.

Seventeen English listeners participated in the ABC condition at Georgetown University, and 11 Guébie listeners participated in the ABC condition in Gnagbodougnoa, Cote d'Ivoire. These participants were different from those who completed Experiment 1.

In the ABC condition, participants were again presented with three nonce words in each trial and asked to decide if they found the third word more similar to the first word in the sequence or

the second word in the sequence. However, in this condition, the third word, C, was never the same type as the A or B word. For example, a speaker could hear [ada] as sound A and [ala] as sound B, followed by [aɗa] as sound C: [ada]$_A$ [ala]$_B$ [aɗa]$_C$. Speakers were asked to choose whether sound C was more similar to sound A or B.

The stimuli and procedure for Experiment 2 are the same as Experiment 1. The experiment started with a practice phase containing 12 trials. After the practice phase, the testing phase began, which contained 100 trials. Given the results of Experiment 1, which show that listeners are not likely to make errors across the place of articulation, the ABC condition was shortened so each trial only compared across the same place of articulation. The experiment was additionally shortened to include only those trials that contained an implosive in at least one of the A, B, or C positions. Had all combinations of all 11 stimuli in each of the A, B, and C positions been used, there would have been a total of 990 trials, which would have taken more than an hour and a half to complete. As participants reported that even the 20-min Experiment 1 felt long, we chose to shorten the ABC trial as much as possible. The remaining 100 combinations of three nonce words were presented in a counterbalanced order.

4.2 Experiment 3 methods

Experiment 3, referred to here as the Mixed condition, was included to see if condition type affects participant responses. In the ABX trials, there is always a "correct" answer. In the ABC trials, participants do not have the opportunity to have a "correct" response. Including the Mixed condition allows us to compare how participants respond to the ABC trials in the task when they are not told whether there is a "correct" answer for each individual trial.

One listener group completed Experiment 3. Seventeen English listeners, who were different from those who completed Experiments 1 and 2, participated in the Mixed condition at Georgetown University. The results from these listeners are discussed below.

The Mixed condition contained all of the ABC trials in addition to either the ABB/ABA or BAA/BAB set of ABX trials. The trials were randomized in PsychoPy. The stimuli and procedure for Experiment 3 are the same as Experiments 1 and 2. Like the ABC condition, both the ABC and ABX trials in the Mixed condition only included trials that compared across the same place of articulation. Again, this was motivated by results from the ABX condition that show participants were not likely to make errors across the place of articulation, and allowed Experiment 3 to be of reasonable length. The Mixed condition thus contained 172 trials in the testing phase.

The analysis of Experiment 3 includes a descriptive comparison of the results of the ABC portion of Experiment 3 compared with Experiment 2. Because of the limited resources available to running perception experiments with Guébie listeners, only English listeners completed Experiment 3. Due to this limiting factor, descriptive results will be compared for English listeners across Experiment 2 and Experiment 3 to determine whether task type impacts the results found.

4.3 Experiment 2 and 3 results

For Experiment 2 and the ABC portion of Experiment 3, Observed/Expected values were calculated (Pierrehumbert, 1993). The O/E value compares the observed proportion of responses for each sound to chance. The Expected value in O/E measures is how many times a result is predicted to occur at chance, whereas the Observed value is the number of times a result is actually observed. For the present study, the Expected value is the number of times a participant is expected to choose a consonant as similar to another consonant if this decision is random (chance). The Observed value is the actual number of times a participant chose a given consonant as most similar to an

Table 5. English Listener Experiment 2 Results: O/E Values Calculated for Sound C = aɓa Compared with Each Consonant in A/B Position.

Sound A or B	Expected	Observed	O/E
aba	85/170	149/170	1.75
apa	85/170	118/170	1.39
ama	85/170	90/170	1.06
ala	85/170	60/170	0.706
aja	85/170	46/170	0.541
awa	85/170	47/170	0.553

Table 6. English Listener Experiment 2 Results: O/E Values Calculated for Sound C = aɗa Compared with Each Consonant in A/B Position.

Sound A or B	Expected	Observed	O/E
ada	85/170	151/170	1.78
ata	85/170	115/170	1.35
ana	85/170	100/170	1.18
ala	85/170	63/170	0.741
aja	85/170	48/170	0.565
awa	85/170	33/170	0.388

implosive. O/E values include calculations for each ABC trial where an implosive consonant is in the C position. Thus, for instances when [aɓa] is in the C position, expected values are calculated as the number of times each bilabial plosive or approximant was present in either the A or B position divided by two. The observed value is calculated as the number of times each consonant in the A or B position was actually chosen when compared with [aɓa]. The same calculations were performed for [aɗa] in the C position compared with alveolar plosives and approximants. Trials with other consonants in the C position were considered fillers and excluded from the O/E calculations.

Subsequently, two chi-square tests were run for each listener group for the ABC tasks. One chi-square test was run for the O/E results with [aɓa] in the C position, and one chi-square test was run for the O/E results with [aɗa] in the C position for each listener group. The chi-square tests determine whether the observed distribution of sounds chosen as similar to implosives is different from chance.

4.3.1 English listener results. The O/E values for the bilabial implosive consonant in the C position are presented in Table 5. Following Pierrehumbert (1993), O/E values greater than 1 are above chance, and O/E values less than 1 are below chance.

The O/E results show that English listeners are most likely to consider the bilabial implosive consonant as similar to the voiced bilabial plosive, followed by the voiceless bilabial plosive, and the bilabial nasal. Table 5 shows that English listeners are not likely to categorize [aɓa] as an approximant.

O/E results for how English listeners categorize the alveolar implosive consonant in the C position of the ABC task are presented in Table 6.
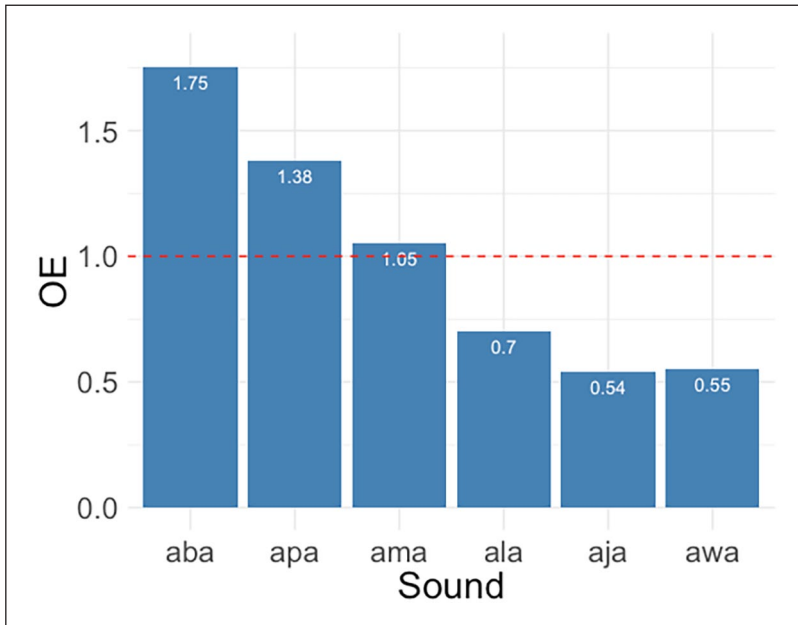
Figure 2. English listener experiment 2 O/E results for Sound C = aɓa.

The alveolar implosive consonant follows the same similarity hierarchy as the bilabial. Participants were most likely to consider the alveolar implosive as similar to the voiced alveolar plosive, followed by the voiceless alveolar plosive, the alveolar nasal, and subsequently, the approximants. To further demonstrate the categorization patterns, Figures 2 and 3 show the similarity ratings for the bilabial implosive compared with bilabial plosives and approximants and the similarity ratings for the alveolar implosive compared with the alveolar plosives and approximants, respectively. The dotted red line represents the expected value if the results were at chance. Sounds with values above the red line are chosen as similar to implosives more often than expected (where "expected" is chance). Sounds with values below the red line are chosen as similar to implosives less often than expected.

Separate chi-square tests were run on the observed results when [aɓa] is in the C position and when [aɗa] is in the C position for the English listener group. The chi-square test when [aɓa] is in the C position shows that each word is not selected with equal frequency ($\chi^2 = 69.12$, $df = 5$, $p = 1.55e-13***$). Similarly, the chi-square test shows that the words are not selected with equal frequency as similar to [aɗa] is the C position ($\chi^2 = 75.01$, $df = 5$, $p = 9.22e-15***$).

Turning now to the results from the English listeners in Experiment 3, the condition type does not seem to impact the consonants that English listeners choose as similar to implosives. O/E values were calculated for the ABC trials of the Mixed condition using the same methods as for Experiment 2. The results for the ABC portion of the Mixed condition are presented in Tables 7 and 8.

When the bilabial implosive is in the C position in the ABC portion of the Mixed condition, English listeners most often choose the voiced bilabial stop as the most similar consonant, followed by the voiceless bilabial stop, and subsequently all sonorants. The alveolar implosive displays the same pattern at the alveolar place of articulation. The voiced alveolar stop is most often chosen as similar to the alveolar implosive, followed by the voiceless stop, and then sonorants.
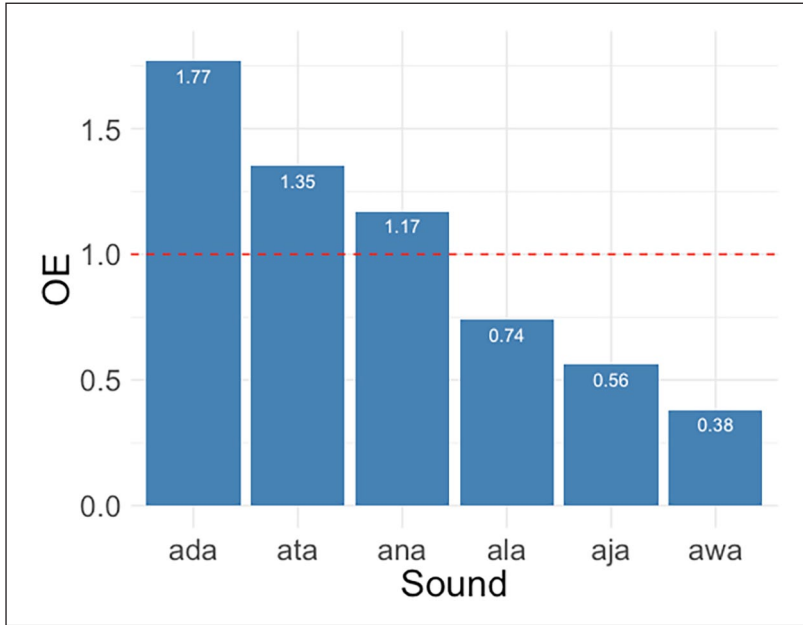
Figure 3. English listener experiment 2 O/E results for Sound C = aɗa.

Table 7. English Listener ABC Portion of Experiment 3 Results: O/E Values for How Often Each Consonant Is Chosen as Similar to Sound C = aɓa.

Sound A or B	Expected	Observed	O/E
aba	85/170	151/170	1.77
apa	85/170	120/170	1.41
ama	85/170	92/170	1.08
ala	85/170	48/170	0.56
aja	85/170	56/170	0.65
awa	85/170	43/170	0.5

Table 8. English Listener ABC Portion of Experiment 3 Results: O/E Values for How Often Each Consonant Is Chosen as Similar to Sound C = aɗa.

Sound A or B	Expected	Observed	O/E
ada	85/170	147/170	1.72
ata	85/170	120/170	1.41
ana	85/170	91/170	1.07
ala	85/170	63/170	0.74
aja	85/170	51/170	0.6
awa	85/170	38/170	0.44

Table 9. Guébie Listener Experiment 2 Results: O/E Values for Each Consonant Compared with Sound C = aɓa.

Sound A or B	Expected	Observed	O/E
aba	60/120	87/120	1.45
apa	60/120	75/120	1.25
ama	60/120	63/120	1.05
ala	60/120	46/120	0.766
aja	60/120	45/120	0.75
awa	60/120	4/120	0.733

Table 10. Guébie Listener Experiment 2 Results: O/E Values for Each Consonant Compared with Sound C = aɗa.

Sound A or B	Expected	Observed	O/E
ada	60/120	73/120	1.22
ata	60/120	67/120	1.11
ana	60/120	59/120	0.983
ala	60/120	54/120	0.9
aja	60/120	52/120	0.866
awa	60/120	55/120	0.916

These results follow the same pattern as the similarity hierarchy resultant from Experiment 2, as in Tables 5 and 6.

There is one minor difference in the results for the ABC portions of Experiment 2 and Experiment 3. English listeners more often chose [aja] as similar to [aɓa] in Experiment 3 than in Experiment 2. This could be due to the results of one participant who often chose [aja] as similar to [aɓa]. However, the general trend holds that English participants are more likely to consider obstruents similar to implosive consonants than sonorants, and the voiced obstruent is the most often chosen. In addition, nasals are chosen as similar to implosives more often than the other sonorants in both the ABC and Mixed conditions at both places of articulation.

4.3.2 Guébie listener results. Guébie listener results for the ABC condition show similar patterns to the English listeners. Looking first at O/E value for the bilabial implosive in the C position, listeners were most likely to categorize [ɓ] as the voiced bilabial stop, followed by the voiceless bilabial stop, and finally sonorants. O/E values for how often each consonant was chosen as similar to the bilabial implosive are given in Table 9.

Results of a chi-square test show that O and E values are significantly different across the consonants chosen as similar to [ɓ] ($\chi^2 = 27.33$, $df = 5$, $p = 4.91\mathrm{e}{-5}$***).

The pattern for which consonants Guébie listeners chose as similar to the alveolar implosive is similar to the English listeners as well. O/E values for how often each consonant was chosen as similar to [aɗa] in the C position are presented in Table 10.

The stimuli containing obstruents, [ada] and [ata], are chosen as similar to [aɗa] slightly more often than chance, while the sonorants are slightly below chance. Note however that the margins by which each observed count differs from the corresponding expected count are smaller than for the bilabial place of articulation (cf. Tables 9 and 10).
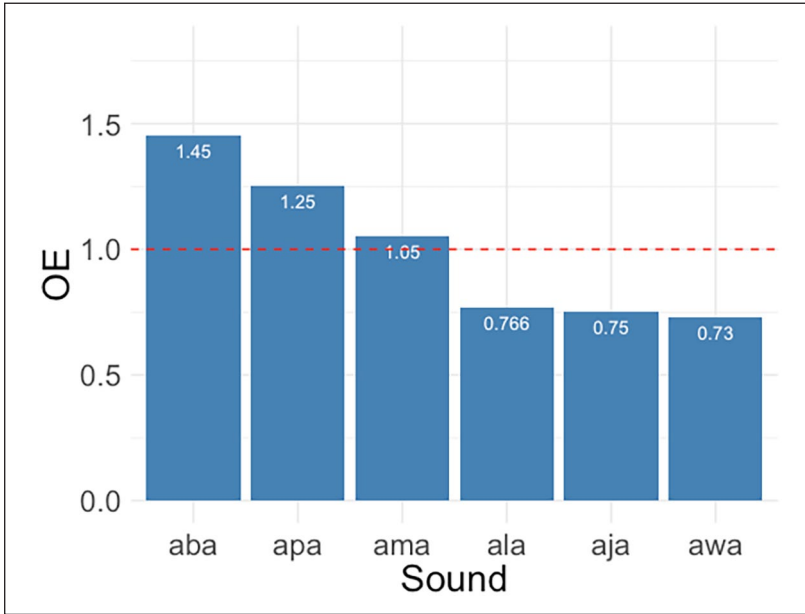
Figure 4. Guébie listeners' experiment 2 O/E results for Sound C = aɓa.
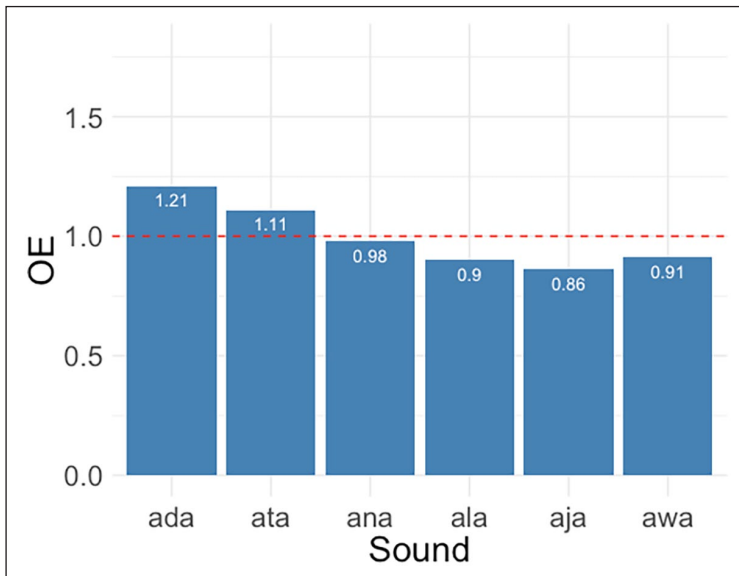

Figure 5. Guébie listeners' experiment 2 O/E results for Sound C = aɗa.

Figures 4 and 5 show how likely each consonant was to be chosen as similar to [aɓa] and [aɗa] by Guébie listeners. Again, the dotted line at one represents chance, the expected values.

Figures 4 and 5 show that across both bilabial and alveolar places of articulation, voiced stops are most often chosen as similar to implosive consonants. However, the O/E values for [aɗa] in the

C position are much closer to chance. A chi-square test reveals that the O/E values are not significantly different across the comparison consonants when [aɗa] is in the C position ($\chi^2 = 5.73$, *df* = 5, *p* = .33). This means that, although alveolar implosives show a similar perceptual pattern to bilabial implosives, [aɗa], unlike [aɓa], is not significantly consistently chosen by Guébie listeners as similar to obstruents over sonorants.

Guébie listeners did not complete Experiment 3. English listeners' results suggest that the ABC portion of the Mixed condition and the ABC task alone produce the same results. Changing the task to underinform participants about whether there was a correct response to each trial in the Mixed condition did not change which consonants English listeners perceived as similar to implosive consonants. In addition, experimental conditions greatly constrained access to eligible Guébie listeners to participate in the study (see the discussion in Section 6). In light of the English listener results that suggest condition type (ABC versus Mixed) does not affect how similar listeners perceive consonants, available Guébie listeners participated in only Experiments 1 and 2.

5 Discussion

5.1 ABX findings: L1 effects on perception

The ABX categorization task results showed a key difference between English- and Guébie-speaking participants. Recall that 73% of errors made by English speakers in the ABX task involved implosive consonants. However, only 53% of errors made by Guébie speakers involved implosives, whereas 51% of the trials included an implosive consonant. The fact that such a high proportion of errors made by English speakers involved an implosive consonant shows that the participants had difficulty categorizing implosive sounds. However, Guébie speakers did not make more errors when an implosive was in the X position than otherwise; their error rates for each sound were proportional to the number of tokens of each sound in the experiment. Furthermore, the output of a mixed-effects logistic regression model for all speaker results shows that there is a significant interaction between language background and trials that contained implosive consonants. Specifically, Guébie listeners were more likely to make correct categorizations on trials that include implosive consonants than English listeners were.

These results suggest that English speakers, but not Guébie speakers, were perceptually miscategorizing implosives. We analyze these results as due to the L1 effect of a contrastive implosive category in Guébie, but not in English. English speakers could not differentiate between implosives and obstruents, and thus regularly miscategorized implosives as obstruents during the task. Following the assumptions of PAM (C. T. Best, 1995; C. T. Best et al., 2001), English listeners' difficulty perceiving the difference between /ɓ/ and /b/, and /ɗ/ and /d/ is likely caused by listeners assimilating /ɓ/ and /ɗ/ to the voiced plosive categories in their L1, and within-category contrast is difficult to perceive (Same Category assimilation). However, Guébie speakers, who have a native implosive category, did not miscategorize implosives more than other sounds (two-category assimilation). This shows that while Guébie listeners considered the implosive consonants to be perceptually similar to voiced plosives (as shown by the results from the ABC task, discussed in Section 5.2), they were not mapping the implosives to plosive categories, but to separate implosive categories. This result is unsurprising, given the fact that Guébie speakers have contrastive voiced plosives and implosives, but importantly shows a baseline for how to assess native speaker perception that has been absent from previous studies. This result further shows that Guébie listeners did not perceptually categorize the non-native implosive stimuli as plosives. Thus, although there is variation in the production of implosive consonants across languages (Ladefoged, 1968), the non-native implosive stimuli in the present study were not perceived as plosives by the Guébie listeners.

The results from Section 3.2 show one unexpected result for the Guébie listeners. Guébie listeners were likely to make categorization errors when comparing the labiovelar glide [w] compared with the bilabial implosive [ɓ]. This result is surprising, given the fact that [w] and [ɓ] are contrastive in Guébie, but shows an important relationship between phonological patterning and perception. First, [w] is produced with the same primary articulators as [ɓ]. Both of these sounds are produced with a bilabial constriction. The shared constriction location likely contributes to the categorization errors of [w], but does not fully explain why no other sound was consistently misperceived as the bilabial implosive (specifically the bilabial plosive and bilabial nasal). Listeners made errors when categorizing glides as implosives, which pattern together phonologically. We hypothesize that the phonological patterning of glides and implosives, together with the constriction location, contribute to the misperception of [w] as [ɓ].

Turning to the alveolar implosive, the fact that Guébie speakers have a bilabial implosive in their L1 resulted in their ability to successfully categorize implosives in the ABX task, even at the alveolar place of articulation. Despite not having a contrastive implosive at the alveolar place of articulation, when tasked with categorizing an alveolar implosive, Guébie speakers performed markedly better than English speakers. Guébie listeners were able to generalize the perceptual contrast between the bilabial implosive /ɓ/ and plosive /b/ to the novel alveolar implosive /ɗ/ and alveolar plosive /d/, correctly categorizing [ɗ] when compared with [d]. We discuss how two theoretical accounts may explain why we see this result for Guébie listeners.

First, as discussed in Section 1.2 and following a distinctive feature theory approach, implosives have been argued to have laryngeal features which distinguish them from plosives (Hall, 2007; Halle & Stevens, 1971; Keating, 1984; Lombardi, 1995). Specifically, Keating (1984) and Lombardi (1995) propose that implosives have a [constricted glottis] feature that distinguishes implosives from voiced plosives. If we assume that Guébie has a similar feature that distinguishes implosives from plosives, then the results here show that listeners have access to this featural information and are extending the [CONSTRICTED GLOTTIS] feature to the alveolar implosive, thus allowing Guébie listeners to correctly categorize the alveolar implosive stimuli as /ɗ/ and not /d/ in the experiments.

Alternatively, PAM (C. T. Best et al., 2001) and articulatory phonology (Browman & Goldstein, 1992) claim that the articulatory gesture is the basic phonological unit, and, important for the present discussion, the object of speech perception. Although /ɓ/ and /b/ share a primary constriction location, C. C. Best and McRoberts (2003) and C. T. Best et al. (2016) define the difference between /ɓ/ and /b/ as a difference in the $+/-$LARYNX-LOWERING GESTURE. Under this perspective, Guébie listeners are able to analogically extend this gesture to the alveolar consonant contrast between /ɗ/ and /d/.

Although the distinctive feature approach and articulatory gesture approach both correctly explain the perceptual patterns observed in the ABX task, PAM and articulatory phonology importantly also make explicit predictions about which contrasts will be difficult for non-native speakers to correctly categorize. Namely, it is predicted that non-native listeners will have difficulty perceiving contrasts that are made using the same primary organ (C. T. Best et al., 2016), which can be interpreted as difficulty perceiving the contrast between non-native sounds that are made using the same primary constriction location. As shown in Section 3.2, listeners are more likely to make categorization errors in trials that contain consonants at the same place of articulation. Guébie listeners do tend to miscategorize [w] as [ɓ], and these two sounds share a constriction location. Gestural models of speech perception predict that listeners will make less categorization errors across the place of articulation, which is what we see with the non-native listeners in this study (both Guébie and English listeners). The relationship between articulatory and perceptual similarity will be further discussed in Section 5.2.

5.2 ABC findings: a cross-linguistic implosive perceptual similarity hierarchy

While we saw L1 effects on the categorization of implosive sounds in the ABX task, we did not see any differences in results from English versus Guébie speakers in the forced-choice similarity ABC task. The results from Experiments 2 and 3, the ABC and Mixed conditions, show that when forced to choose which of two consonants is most similar to an implosive, English and Guébie speakers make similar judgments. Both English and Guébie speakers chose voiced plosives as most similar to implosives, followed by voiceless plosives, then nasals, then approximants. The results from this experiment yield a hierarchy of consonants perceived as similar to implosives [ɓ] and [ɗ], (5). Because we found the same results in both groups, despite differences in L1 exposure to contrastive implosives, and despite the fact that implosives pattern phonologically with sonorants in Guébie but are not perceived as similar to sonorants, we hypothesize that this is a universal perceptual similarity hierarchy. Although not feasible to run the same experiment with speakers of a language where implosives are contrastive and pattern with obstruents, we expect the same hierarchy would emerge based on the current findings that both English and Guébie listeners considered the plosives at the same place of articulation to be most perceptually similar to implosives.

5. Hierarchy of consonant similarity to implosives

        ɓ: b ≫ p ≫ m ≫ l/j/w

        ɗ: d ≫ t ≫ n ≫ l/j/w

Due to the slight variation across speaker groups and experiment conditions in which approximants were chosen as most similar to implosives, we have not ranked approximants with respect to each other in the hierarchies in (5). The sound [w] was chosen more often as similar to [ɗ] by Guébie speakers than either [l] or [j], which differs slightly from the English speaker results. However, the difference between the frequency that each of these three sounds was chosen by Guébie speakers is so small that it is likely due to chance. The overall generalization that obstruents are chosen more often than sonorants holds across languages.

Summarizing the ABC results, when forced to choose, both English and Guébie speakers perceive implosives as most similar to obstruents, not sonorants. The resulting hierarchy is not affected by the presence or absence of an implosive in a speaker's L1. These results may seem surprising for Guébie speakers, who have a contrastive implosive /ɓ/ in their L1 inventory that patterns phonologically with sonorants, and given the fact that Guébie listeners often misperceive [w] as [ɓ]. The phonological patterning of /ɓ/ with sonorants seems not to affect or be affected by the cross-linguistic perceptual similarity of implosive [ɓ] to plosives of the same place of articulation.

This finding shows that articulatory similarity, defined by the primary organs involved in making a sound contrast, is more predictive of how similar two sounds will be perceived than L1 phonological patterning is. As discussed above, [ɓ] and [b] share a primary constriction location and degree, and [d] and [ɗ] share a primary constriction location and degree. These sounds differ in laryngeal configuration. Both Guébie listeners and English listeners classified the implosives and obstruents as similar, showing that the primary organs involved in making a contrast are the most important factor in whether two sounds are considered similar, regardless of the phonological patterning of the sounds in a given language.

Further strengthening this claim are the results showing which sounds were chosen as the next most similar to implosives, after voiced plosives. For both the Guébie and English listeners, the voiceless plosives [p] and [t], followed by the nasals [m] and [n] were next most frequently selected as similar to implosives in the ABC task. [p] and [m] also share a constriction degree and location with the implosive [ɓ], and [t] and [n] share a constriction degree and location with [ɗ]. For the

approximants, [j] is produced with a constriction location at the palate rather than the alveolar region or with the upper and lower lips. [l] is produced with a complete apical closure, but importantly is also characterized by dorsal gestures (see Sproat and Fujimura (1993) for a discussion of the articulatory gestures of laterals in English). Despite the fact that [w] and [ɓ] share constriction location and Guébie listeners misperceived [w] as [ɓ] in the ABX task, neither Guébie listeners nor English listeners perceived [w] as more similar to [ɓ] than they did [b] or [p]. This could be because [w] has an additional velar constriction, but also could be because [w] does not share a constriction degree with [ɓ]. All the approximants, in fact, are produced with a wider opening at the primary constriction location than plosives, implosives, and nasals. Thus, the hierarchy of sounds chosen as perceptually similar to implosives appears to be motivated by the shared constriction degree and primary constriction location of the two sounds.

A note must be made here about the alveolar implosive results for Guébie listeners. Despite the fact that both listener groups followed a consistent hierarchy of sound similarity in the ABC task, the distribution of sounds chosen as similar to [aɗa] by Guébie speakers did not reach significance. This suggests that Guébie listeners were less consistent in their responses for the ABC task when the alveolar implosive was presented than the bilabial implosive. Guébie listeners have a bilabial implosive category in their L1, and not an alveolar implosive, which could be responsible for the differences in responses for the two implosives. However, the results for English speakers, who have familiarity with neither a bilabial nor alveolar implosive, came out as significant for both sounds. Overall, the Guébie listeners show a smaller difference from expected values for both the alveolar implosive and the bilabial implosive, as can be seen in Figures 4 and 5 compared with English listener results, as shown in Figures 2 and 3. This may be a task effect, as overall Guébie listeners were much less familiar with this type of task, and therefore were more likely to have inconsistently selected a response. The limitations associated with running perception tasks in a field-setting are discussed further in Section 6. Despite a lack of significance for the alveolar implosive for Guébie speakers when compared with chance, the raw hierarchy of sound similarity in (5) held across both implosive categories for English and Guébie listeners.

The ABC task is used here to discover a hierarchy of perceptual consonant similarity. We propose that this tool be adopted more broadly as a mechanism for examining sound similarity within and across languages, much like a confusion matrix. This task has the added benefit of avoiding orthography in perceptual similarity tasks, which allows for a more diverse range of languages and speaker backgrounds' to be included in perception tasks.

6 Limitations

Our perception results rely on comparing data from English-speaking and Guébie-speaking participants. However, due to factors outside of our control, the two sets of results may not be directly comparable. This section discusses these potential limitations.

First, there is a possibility that the selection of Laal and Vietnamese speakers to produce the implosive stimuli affected the perception results. Implosives show different phonological patterns in Vietnamese and Laal, which could be related to different phonetic properties of implosives in these two languages. In Laal, implosives pattern as sonorants, as evidenced by implosives nasalizing word initially (as sonorants do), and an alternation between [ɗ] and [r] (Sande & Oakley, 2021, Florian Lionnet, p.c.). The patterning of implosives as sonorants or obstruents is less clear in Vietnamese because of the phonological inventory of the language. Vietnamese does not have contrastive voiced stops and implosives, which makes it difficult to determine whether implosives pattern phonologically with voiced stops (Greenberg, 1970, p. 137). However, there are two main reasons to suspect that the language background of the speakers who produced the stimuli did not greatly impact the results here. First, both speakers were trained linguists who were using their

knowledge of phonetics as well as their language background to produce the implosives. Second (and most importantly), the results from the ABX task show that the L1 Guébie listeners correctly categorized the implosives as implosives, whereas the L1 English listeners often categorized implosives as voiced obstruents. The Guébie listeners were able to categorize implosives stimuli produced by speakers of languages besides Guébie more accurately than English listeners, presumably because this listener group has an implosive in their L1, despite the fact that the implosives were not produced by a Guébie speaker. Furthermore, if the language background of the speakers who produced the stimuli were to affect the results, we might expect Guébie listeners to be more likely to perceive implosives as similar to sonorants because implosives are sonorant patterning in Laal. However, the ABC task results show the opposite pattern. The cross-linguistic production of implosives and the phonological patterning of implosives warrants much more research, but for the reasons discussed here, it is unlikely that the language background of the speakers who produced the stimuli affected the results of the present study.

Second, although both English and Guébie speakers performed better than chance during the ABX task, Guébie speakers made many more errors than English speakers overall. We hypothesize that the lower performance of Guébie speakers is due to a lack of familiarity with the type of task and with using computers. The Guébie listeners all grew up in Gnagbodougnoa, Côte d'Ivoire, and were living there at the time of their participation in the experiment. Gnagbodougnoa is a village of 1,000, situated in the jungle of southwest Côte d'Ivoire. Until 2009, there was no high school nearby. Most of the population of Gnagobdougnoa are subsistence farmers, who are not often faced with metalinguistic tasks, and who have limited access to education and technology. The English listeners, however, were undergraduate students at Georgetown University who have been exposed to test-taking and language courses for all of their lives.

As summarized here and in the methodology, there were a number of challenges associated with running a perceptual experiment in PsychoPy in the Guébie community. We expect that similar challenges hold in fieldwork settings in developing countries in general, particularly in rural areas where the population has little access to education or technology. These challenges may explain the lack of experimental data on minority languages in the phonetic and psycholinguistic literature. We hope that by sharing the challenges faced in running perceptual experiments in Gnagbodougnoa, future researchers running experiments in a fieldwork setting may be better prepared. For researchers who find themselves in similar situations, we recommend keeping the task and technology interface as simple as possible, and collaborating with a native speaker who can provide instructions to participants in the language of the community.

7 Conclusions and directions for future work

To conclude, we found that speakers with L1 implosive contrasts correctly categorize implosives as distinct from plosives at the same place of articulation, and we confirmed that English speakers, who have no contrastive L1 implosive sound, often miscategorize implosives. Listeners who have one implosive category in their L1 were accurate at categorizing non-native implosives at other places of articulation, showing that listeners can generalize L1 articulatory gestures or features to non-native segments in speech perception. In addition, we found that despite the fact that implosives pattern phonologically with sonorants to the exclusion of obstruents in Guébie, a minority language of Côte d'Ivoire, Guébie listeners perceive implosive sounds as more similar to obstruents than sonorants, just as naïve listeners do. This suggests that, as predicted by PAM, shared major articulators better predict perceptual similarity than does L1 phonological patterning.

In addition to our findings on perceptual similarity, we have shown that an ABX-style task where the third sound X is different from the first two sounds, A and B (our so-called ABC task in

Section 4), can serve as a tool for investigating sound similarity in languages where speakers may not be literate or a language may not have an orthography.

ORCID iDs

Madeleine Oakley https://orcid.org/0000-0002-6459-0658
Hannah Sande https://orcid.org/0000-0003-1335-8717

References

Bennett, W. G. (2013). *Dissimilation, consonant harmony, and surface correspondence* [PhD thesis]. Rutgers University.

Best, C. C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, *46*(2–3), 183–216.

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. 171–206). New York Press.

Best, C. T., Goldstein, L. M., Nam, H., & Tyler, M. D. (2016). Articulating what infants attune to in native speech. *Ecological Psychology*, *28*(4), 216–261.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775–794.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech. *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, *17*, 13–34.

Bodji, S., & Sande, H. (2014). *Guébie fieldwork collection* (Bundle 2014-15.123, California Language Archive, Survey of California and Other Indian Languages). University of California, Berkeley.

Bohn, O.-S., & Best, C. T. (2012). Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *Journal of Phonetics*, *40*(1), 109–128.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3–4), 155–180.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Harper and Row.

Clements, G. N., & Hume, E. V. (1995). The internal organization of speech sounds. *The Handbook of Phonological Theory*, *7*, 245–317.

Clements, G. N., & Osu, S. (2002). Explosives, implosives and nonexplosives: the linguistic function of air pressure differences in stops. In C. Gussenhoven (Ed.), *Laboratory phonology 7* (pp. 299–350). Mouton de Gruyter.

Faris, M. M., Best, C. T., & Tyler, M. D. (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *The Journal of the Acoustical Society of America*, *139*(1), EL1–EL5.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, *92*, 233–277.

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 University Press.

Greenberg, J. H. (1970). Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics*, *36*(2), 123–145.

Hall, T. A. (2007). Segmental features. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 311–334). Cambridge University Press.

Halle, M., & Stevens, K. N. (1971). A note on laryngeal features. In M. Halle & K. N. Stevens (Eds.), *From memory to speech and back* (pp. 45–61). De Gruyter Mouton.

Hansson, G. (2001). The phonologization of production constraints: Evidence from consonant harmony. *Chicago Linguistic Society*, *37*, Article 187.

Hume, E., & Johnson, K. (2001). *The role of speech perception in phonology*. Academic Press.

Jakobson, R., Fant, C. G., & Halle, M. (1951). *Preliminaries to speech analysis: The distinctive features and their correlates*. MIT Press.

Kaye, J. (1981). Implosives as liquids. *Studies in African Linguistics*, *1981*(Suppl. 8), 878–881.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60*, 286–319.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000.

Ladefoged, P. (1968). *A phonetic study of West African languages: An auditory-instrumental survey*. Cambridge University Press.

Lombardi, L. (1995). *Laryngeal Features and Privativity*. Walter de Gruyter.

Lüdecke, D. (2021). *sjPlot: Data visualization for statistics in social science* (R package version 2.8.10). R Foundation for Statistical Computing.

Marchese, L. (1979). *Atlas Linguistique Kru*. International Languages Academy.

Mielke, J. (2004). *The emergence of distinctive features* [PhD thesis]. The Ohio State University.

Mielke, J. (2005). Ambivalence and ambiguity in laterals and nasals. *Phonology*, *22*(2), 169–203.

Mielke, J. (2008). *The emergence of distinctive features*. Oxford University Press.

Mielke, J. (2012). A phonetically based metric of sound similarity. *Lingua*, *122*(2), 145–163.

Newman, P. (2000). *The Hausa language: An encyclopedic reference grammar*. Yale University Press.

Pajak, B., & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of Phonetics*, *46*, 147–160.

Paradis, C. (1992). *Lexical phonology and morphology: The nominal classes in Fula*. Garland Publishing.

Pierrehumbert, J. B. (1993). Dissimilarity in the Arabic verbal roots. *Proceedings of the North East Linguistics Society*, *23*, 367–381.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rose, S., & Walker, R. (2004). A typology of consonant agreement as correspondence. *Language*, *80*, 475–531.

Sande, H. (2017). *Distributing morphologically conditioned phonology: Three case studies from Guébie* [PhD thesis]. UC Berkeley.

Sande, H., & Oakley, M. (2021). A typological survey of the phonological behavior of implosives: Implications for feature theories [Manuscript]. UC Berkeley and North Carolina State University.

Sproat, R., & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, *21*(3), 291–311.

Strange, W. (2007). Cross-language phonetic similarity of vowels. *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, *17*, 35–55.

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, *39*(4), 456–466.

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*(4), 311–344.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4–21.

Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*, Article 1000.