



HAL
open science

The perception of vowelless words in Tashlhiyt

Mohamed Lahrouchi, Georgia Zellou, Karim Bensoukas

► **To cite this version:**

Mohamed Lahrouchi, Georgia Zellou, Karim Bensoukas. The perception of vowelless words in Tashlhiyt. *Glossa: a journal of general linguistics* (2021-..), 2024, 8 (1), pp.10.16995/glossa.10438. 10.16995/glossa.10438 . hal-04358503

HAL Id: hal-04358503

<https://hal.science/hal-04358503>

Submitted on 21 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The perception of vowelless words in Tashlhiyt

Zellou, Georgia¹, Lahrouchi, Mohamed² & Bensoukas, Karim³
¹UC Davis, ²CNRS & Université Paris 8, ³Université Mohammed V

Abstract

This study examines the perceptual mechanisms involved in the processing of words without vowels, a lexical form that is common in Tashlhiyt but highly dispreferred cross-linguistically. In Experiment 1, native and naive (English-speaking) listeners completed a paired discrimination task where the middle segment of the different-pair contained either a different vowel (e.g., *fan* vs. *fin*), consonant (e.g., *ʌbr* vs. *ʌdr*), or vowelless vs. voweled contrast (e.g., *tlf* vs. *tuf*). Experiment 2 was a wordlikeness ratings task of Tashlhiyt-like tri-segmental nonwords constructed to vary in the sonority of the middle segment. We find that vowelless words containing different types of sonority profiles are generally discriminable by both native and naive listeners. This can be explained by the phonetic and acoustic properties of vowelless words: Since Tashlhiyt exhibits low consonant-to-consonant coarticulation, the presence of robust consonantal cues in the speech signal means that the internal phonological structure of vowelless words is recoverable by listeners. Moreover, speech style variation provides further evidence that the phonetic implementation of vowelless words makes them perceptually stable. At the same time, wordlikeness ratings of nonwords indicate that listeners rely on their native-language experience to process the wellformedness of new words: Tashlhiyt listeners accept sonorant- and obstruent-centered vowelless words equally; meanwhile, English listeners' preferences increase with higher sonority values of the word center. Thus, our findings provide an overview of the low-level acoustic-phonetic and higher-level phonological processing mechanisms involved in the perception of vowelless words. Our results can inform understandings of the relationship between language-specific phonetic variation and phonotactic patterns, as well as how auditory processing mechanisms shape phonological typology.

Keywords: Vowelless words, Sonority, Perception, Tashlhiyt, Phonetic variation

1. Introduction

Identifying the forces that shape the form of words across languages can inform phonological theory. One fundamental observation about phonology is that certain sequences of speech sounds are favored, and others dispreferred, in the forms of words and syllables found cross-linguistically. For instance, across languages, there is an overwhelming preference for words to contain vowels. However, in some languages, such as Tashlhiyt, words can contain sequences of only consonants. Why are vowelless words so cross-linguistically uncommon? Understanding this asymmetry in word shapes can shed light on some of the fundamental questions about how sound patterns emerge and evolve in linguistic systems. Tashlhiyt is even more unique in that not only are vowelless words permitted but also words with sequences of obstruents are allowed and are common within the language. There is a body of work investigating the articulatory properties of vowelless words in Tashlhiyt, as well as in other languages where they are

permitted, some with a focus on understanding what makes them phonologically stable (Ridouane, 2008; Ridouane & Fougeron, 2011; Fougeron & Ridouane, 2008; Pouplier & Beňuš, 2011). Yet, another theoretical approach is that auditory factors can also provide insight into the stability of a phonological system (Ohala, 1993; Beddor, 2009; Harrington et al., 2019). For instance, some have argued that observed cross-linguistic phonological tendencies are the result of auditory properties of the speech signal or perceptual processing mechanisms (Ohala, 1981; Blevins, 2004). A focus on the perception of vowelless words can provide insight into how they emerge and are maintained in a phonological system.

The focus of the present study is to investigate the perception of words without vowels in Tashlhiyt. We asked whether there are clues in the perceptual mechanisms involved in the processing of vowelless words that can speak to their rare typological distribution. To this end, we examine lexical discrimination performance and nonword acceptability judgements of tri-segmental vowelless words containing different sonority profiles by native Tashlhiyt and non-native (English-speaking) listeners. Across these tasks, we also manipulated speaking style (clear vs. casual speech) to explore the effect of systematic hypo- and hyper-articulatory variation on the perceptual patterns.

1.1. Tashlhiyt and vowelless words

Tashlhiyt (iso: [shi]) is an Amazigh (Berber; Afroasiatic) language spoken in southern Morocco with an estimated 5 million speakers. The phoneme inventory of Tashlhiyt is consonant-heavy: there are 34 consonants (/b, m, f, t, t^h, d, d^h, n, r, r^h, s, s^h, z, z^h, l, l^h, ʃ, ʃ^h, ʒ, ʒ^h, j, k, k^w, g, g^w, w, q, q^w, χ, χ^w, β, β^w, ħ, ʕ, h/), which contrast as singleton and geminate in all word positions, and three vowels (/a, i, u/) (Ridouane, 2014). Tashlhiyt permits highly complex syllable structures, which are more likely to be found in languages with higher consonant-to-vowel inventories (Maddieson, 2013; Easterday, 2019). Phonotactics in Tashlhiyt are extremely permissive: words containing consonant sequences that challenge cross-linguistic sonority tendencies are common, for instance, word-initial sequences containing plateau /kti/ or falling /rku/ sonority profiles (Dell & Elmedlaoui, 2002; Ridouane et al., 2014; Lahrouchi, 2010, 2018; Jebbour, 1996, 1999; Boukous, 1987, 2009). Another cross-linguistically rare pattern in Tashlhiyt is the presence of “vowelless” words containing only consonants and no lexical vowels, e.g., /tftktstt/ ‘you sprained it’ (Dell & Elmedlouai, 2002; Ridouane, 2008; though other varieties of Amazigh do display schwa epenthesis).

A very robust cross-language phonological tendency is for words to contain vowels. This is consistent with phonotactic sequencing accounts that syllables containing rises in sonority from periphery to nucleus are preferred (Clements, 1990; Zec, 1995). Yet, it has also been shown that constraints on sonority are language-specific and vowels-as-nuclei is not an absolute; for instance, some languages allow consonants to be syllabic. Even within the languages that allow consonant nuclei, sonorant consonants are the most commonly permitted syllabic segment (Bell, 1970), and in many languages syllabic consonants are restricted to some environments (e.g., in unstressed syllables in German and English, e.g., *bottle*, *button*). Allowance of consonant nuclei in stressed positions (e.g., monosyllabic words) is less common, but can be found in languages such as Slovak (syllabic liquids) and Yoruba (syllabic nasals). The most cross-linguistically rare phenomenon is that of syllabic obstruents.

Tashlhiyt allows any segment to occupy the syllable nucleus (Dell & Elmedlaoui, 1985, 2002; Bensoukas, 2001; Ridouane, 2008; Ridouane, 2014; Lahrouchi, 2018). Moreover,

vowelless words are frequent in Tashlhiyt; for instance, Ridouane (2008: 328) reports that in a collection of texts, 22% of the lexical items are vowelless words and around 8% of words contain voiceless obstruents only. Lahrouchi (2010) compiled a list of 222 bi- and tri-segmental Tashlhiyt verbs and categorized them into classes based on sonority profile of C1-C2 (note that a focus on word-onset sequences is consistent with work showing that the coordination of C1 and C2 in vowelless words is more constrained than that for C2 and C3, Pouplier & Beňuš, 2011): about 20% of the verbs in his database are vowelless words containing an initial sonorant-obstruent sequence (rising sonority, e.g., *krz* ‘plow’), 17% contain an initial cluster of obstruents (plateauing sonority; e.g., *gzm* ‘cut’), 15% contain an initial obstruent-sonorant sequence (falling sonority; e.g., *rgl* ‘knock’). 33% contain vowels (e.g. *knu* ‘lean’) and just 2% contain plateauing sonority with initial sonorants (SSO or SSS). Thus, within tri-segmental vowelless words in Tashlhiyt, rising (OSC), falling (SOC), and plateauing (OOC) sonority profiles are all frequent forms. (Note that the morphology of the language is sensitive to the syllable structure of consonant-only verbs. A case in point is imperfective internal vs. external gemination, as in the monosyllabic verb *krz/kkrz* ‘plow’ and disyllabic verb *mgr/mggr* ‘harvest’. (Dell & Elmedlaoui, 1985; Jebbour, 1996; Bensoukas, 2001).)

There is also a large body of work investigating the articulatory and acoustic properties of vowelless words in Tashlhiyt. Ridouane (2008; Ridouane et al., 2014) uses phonetic and phonological evidence to demonstrate that consonant sequences in Tashlhiyt are heterosyllabic and do not form clusters. In a “complex cluster” language, like English, consonant sequences form temporally overlapping, coarticulated gestures; for instance, in English, the duration of a consonant compresses when onsets contain more than one segment (e.g., the /k/ in *scab* is shorter than that in *cab*, Marin & Pouplier, 2010). However, in Tashlhiyt, as the number of consonants within a word increases, the segment durations remain stable indicating non-overlapping coordination of segments (Ridouane et al., 2014; Hermes et al., 2011; Hermes et al., 2017). Tashlhiyt listeners also treat consonant sequences heterosyllabic: while participating in a game where they heard a lexical item containing an initial consonant sequence and then had to repeat back only the “first part”, Tashlhiyt speakers responded overwhelmingly with simpleton onsets, indicating that they parse consonant sequences into multiple independent syllables (Ridouane et al., 2014).

Furthermore, Tashlhiyt has been shown to have “wide” timing of sequential consonants (Fougeron & Ridouane, 2008). In fact, in many languages that allow words with highly complex consonant sequences, the gestures for sequential consonants are timed far apart from one another (Gafos, 2002 for Moroccan Arabic; Pouplier & Beňuš, 2011 for Slovak; Tilsen et al., 2012 for Hebrew). This “non-coarticulating consonants” property also appears to be a common feature of languages that permit vowelless words (Pouplier & Beňuš, 2011). This often results in an intrusive acoustic element between consonants, ranging in quality from a vocoid to a full schwa-like element (Kirby, 2014; Hall 2006). In Tashlhiyt, the vocoid or schwa does not increase the length of the sequence when it occurs (Ridouane & Fougeron, 2011), compared to epenthesis in other languages where the addition of a schwa increases word duration (Davidson & Roon, 2008; Hall, 2006). In Tashlhiyt, the presence of these variable vocoid elements in consonant sequences is argued to reflect a transitional acoustic signal due to production of consonants that are coordinated to be timed far apart from one another, rather than an epenthetic vowel (Dell & Elmedlaoui, 2002; Ridouane, 2008; though, cf, Coleman, 2001).

1.2. Perceptual explanations for the rarity of vowelless words

The existence, and prevalence, of vowelless words in Tashlhiyt presents an opportunity to address a fundamental puzzle in phonological theory: why are vowelless words so rare across languages of the world? In the current study, we focus on possible perceptual explanations. Some argue that cross-linguistic phonological tendencies can be explained by perceptual processing mechanisms since many common phonological patterns have perceptual motivations (Ohala, 1993; Blevins, 2004; Beddor, 2009; Harrington et al., 2019). Perception biases could be at play in the cross-linguistic preference for CV structure: it is harder to identify a consonant when it appears in a cluster (e.g., in English, listeners recognize the “b” in *band* faster than the “b” in *brand* or *bland*, Cutler et al., 1987); and a speech stream containing CV sequences will contain maximally distinct acoustic modulations which are more salient, and thus recoverable, to a listener than a speech stream containing only consonants or vowels (Ohala & Kawasaki-Fukumori, 1997).

Our approach follows theoretical frameworks exploring how the mechanisms of human speech processing can be a source of understanding the stability of sound patterns and how they emerge and evolve over time (Ohala, 1993; Blevins, 2004; Harrington et al., 2019). By investigating perceptual patterns of synchronic speech variants using laboratory methods, phonological forms that are harder for listeners to recover can be identified. As outlined above, there is much work examining the gestural dynamics of vowelless words in Tashlhiyt (e.g., Ridouane et al., 2014). Yet, investigating the *perception* of vowelless words is comparatively under-researched. We ask whether examining the perception of vowelless words in Tashlhiyt can provide clues as to why they are so cross-linguistically rare, and also perhaps, the phonetic pre-conditions that permit them to be maintained in Tashlhiyt.

One of the auditory mechanisms that might make vowelless words susceptible to mistransmission is that consonantal cues can be obscured when surrounded by other consonants (Wright, 1996, 2004). Perceptual cues vary based on the neighboring sounds. This means that the “same sound” can be more or less perceptually recoverable depending on the context in which it occurs. For instance, stops generate robust formant transitions on adjacent vowels, due to the large amount of movement from a closed to an open oral constriction. Formant transitions are an acoustic cue that listeners rely on to identify both the place and voicing of plosives (Liberman et al., 1954; Benkí, 2001). A stop produced before another stop will result in fewer and less robust acoustic cues in the speech signal: A release burst between the first and second consonants would provide the only cues for listeners about the identity of the initial segment; reduction, or loss, of the release burst could jeopardize listeners’ ability to recover the speaker’s intended form (Wright, 1996). Thus, a potential perceptual bias against vowelless words comes from the role of acoustic cues in the maintenance of phonological contrast. We focus on the perception of Tashlhiyt tri-segmental vowelless minimal pairs contrasting in the middle segment.

Moreover, the perceptibility of vowelless words might vary based on sonority patterns. One view is that greater acoustic contrast between adjacent segments allows for more robust transmission from speakers to hearers (Ohala & Kawasaki-Fukumori, 1997). If adjacent segments are *acoustically similar*, they are more likely to be confused by a listener, thus diachronically unstable. Ohala and Kawasaki-Fukumori predict that sound sequences containing plateau sonority patterns (e.g., /sfs/) are the hardest to perceive since they contain the smallest differences in acoustic modulations across segments. Meanwhile, sequences containing larger acoustic perturbations (like, /ble/ or /ske/) are less susceptible to perceptual confusion, thus more

phonologically stable. Evidence supporting this comes from Chen et al. (2022), who investigated Mandarin listeners' discrimination of rising (e.g., *kl*), falling (e.g., *lk*), and plateau (e.g., *kp*) onset clusters with and without vowel epenthesis. Discrimination performance was higher for both rising and falling sonority clusters because the acoustic difference between the consonants makes them more perceptible to listeners, compared to plateauing clusters. For the perception of vowelless words, an *acoustic similarity* hypothesis is that since sequences of consonants are more acoustically similar, vowelless words will be overall harder to discriminate than words with vowels. Moreover, within vowelless words, the prediction is that forms with more similar sonority patterns should be even harder to discriminate. If these predictions are borne out, it could explain why vowelless words would not be robust to transmission and, thus, evolutionarily dispreferred.

On the other hand, as outlined above, prior work on vowelless words in Tashlhiyt finds that they are often produced with vocoid elements between consonants (Ridouane & Fougeron, 2011). Since Tashlhiyt exhibits low consonant-to-consonant coarticulation, the presence of robust cues in the speech signal means that the internal phonological structure of vowelless words could be recoverable by listeners. Some cross-linguistic evidence in support of this possibility comes from Wright's (1996) study of stop+stop consonant clusters in Tsou across different word positions. He found that stop release bursts are absent intervocally, but distinct word-initially where the formant cues are absent. He argued that speakers are actively enhancing secondary acoustic cues for consonants in contexts where they are more likely to be difficult for the listener to perceive; and indeed, when the release bursts are present in initial clusters, listeners accurately identify the stop+stop clusters. Thus, a *cue preservation hypothesis* is that the phonetic implementation of vowelless words in Tashlhiyt creates salient acoustic cues that make them perceptually stable.

It is also possible that perceptual patterns follow predictions made by a traditional *sonority sequencing* account (Clements, 1990). One common way of assessing the perceptual preferences of consonant clusters is to present nonwords to participants (either auditorily, or through orthography) and have them rate the likelihood each item could be a possible word in their language. Studies by Berent and colleagues (e.g., Berent et al., 2009) have demonstrated that speakers from various language backgrounds prefer rising sonority consonant sequences over plateauing sequences, which are in turn preferred over falling sonority profiles (e.g., *blif* > *bnif* > *bdif* > *lbif*; rising > plateau > falling). This is consistent with theoretical accounts that the sonority hierarchy governs sound sequences in an active way in synchronic grammars (Zec, 1995).

1.3. Phonetic variation across speaking styles

Above and beyond their perception more generally, we also ask the impact on listeners when speakers explicitly *enhance* vowelless words. Talkers adapt their speech based on the communicative context: speakers vary between hyper- and hypo-speech variants depending on whether there is evidence that the listener will misunderstand them (Lindblom, 1990). Hyperarticulated speech contains a variety of acoustic enhancements relative to non-clarity-oriented speech, such as longer and more extreme segment realizations (Picheny et al., 1986; Krause & Braidá, 2004; Cohn et al., 2022). Also, moreover, the acoustic effects of clear speech have been shown to increase intelligibility for listeners (Picheny et al., 1985; Smiljanić & Bradlow, 2011). While clear speech effects have been found across several languages (e.g.,

Smiljanić & Bradlow, 2005 for Croatian, Tupper et al., 2021 for Mandarin, Kang & Guion, 2008 for Korean), more work is needed to understand how enhancement affects different types of phonological contrasts.

Prior work on clear speech reports acoustic enhancements of both consonants and vowels (Krause & Braida, 2004). Therefore, clear speech should improve the perceptibility of all words in Tashlhiyt. However, examining Tashlhiyt enhancement effects can further inform different hypotheses about perception of different vowelless word types. From an *acoustic similarity* account, words containing the largest acoustic differences across segments (i.e., words vowel or sonorant centers) could receive the largest clear speech boost since phonetic enhancement might make those sounds even more acoustically distinct. Evidence for this comes from a recent study of the perception of onset singleton-geminate-cluster contrasts in Tashlhiyt (Zellou, Lahrouchi, & Bensoukas, 2022). They found that discrimination of the rarer onset contrasts (geminate vs. non-rising onset consonant sequences) is harder for listeners in both clear and causal speech; in fact, there was no clear speech boost for contrasts involving the non-rising clusters. We predict similar patterns for vowelless words in the present study.

A *cue enhancement* prediction is that clear speech enhancements are targeted for phonological contrasts that might be particularly challenging for listeners. Wright's (1996) observation of consistent release bursts in stop+stop clusters in contexts where other cues are not present in Tsou is consistent with this. There is also evidence that talkers selectively enhance phonetic cues in response to a listener misunderstanding a target word with a minimal pair competitor (e.g., Seyfarth et al., 2016; Buz et al., 2016). For the present study, if vowelless words containing plateau sonority are the hardest to perceive in reduced speech because the cues to segments are not salient, they could receive the largest perceptual boost in clear speech where the speaker is aiming to make words maximally intelligible to listeners.

Regardless of which hypothesis is supported, the mere existence of alternative forms of words due to hypo- and hyper-articulation provides a rich testing ground to examine the mechanisms of perceptually-based variation (Blevins, 2004). Since vowelless words are typologically rare, we believe that a comprehensive investigation into their perceptual processing should also include systematic acoustic variants since listeners encounter multiple phonetic forms of words. Moreover, descriptively, examining how speech variation affects the range of different linguistic contrasts found across the world's languages is important to understand the relationship between speech communication and phonology. Thus, we examine speech style variation as an additional factor in this present study.

1.4. Non-native speech perception

How do naive listeners perceive vowelless words? There is much work examining how non-native listeners process unfamiliar consonant sequences. For instance, there is evidence that speech processing is sensitive to the sound sequencing probabilities in one's native language (e.g., van der Lugt, 2001; Best et al., 2001; Davidson, 2011). Investigating cross-language speech perception of typologically rare contrasts adds to our scientific understanding about universal processing mechanisms of different types of speech sounds (Bohn, 2017). For instance, cross-linguistically rarer phonotactic patterns are less likely to be rated as a possible word in one's native language (Berent et al., 2009).

Comparing across listeners with different language backgrounds can also inform about the fundamental mechanisms of speech processing. For instance, some researchers have argued

that a near-universal in word identification is the *possible-word constraint* that listeners exploit transitional probabilities when segmenting sequences of sounds into words. For instance, since most languages only have words that contain vowels, listeners from a range of different language backgrounds segment words more easily from a context where the remainder would itself create a word with a vowel (e.g., *apple* from *vuffapple*; *vuff* is an acceptable word form), compared to where the remainder would create a vowelless word (e.g., *fapple*; *f) (Norris et al., 1997). However, El Aissati et al. (2012) found that native Tarifiyt (an Amazigh language related to Tashlhiyt) listeners *do not* exhibit the possible-word constraint when segmenting words in running speech. They argue that fundamental speech processing mechanisms, such as the possible-word constraint, are simply not active during spoken word comprehension by native speakers of a language that allows vowelless words. We explore this possibility further by comparing native and non-native Tashlhiyt listeners' perception of vowelless words in two additional types of tasks.

At the same time, there is work showing that clear speech increases the intelligibility of non-native sound contrasts (Kabak & Maniwa, 2007; Zellou et al., 2022). Since L2 adult learners usually begin as naive listeners, exploring the perception of novel sound contrasts in a foreign language also has applications for second language acquisition (Bohn, 2017). Looking at cross-language perception of vowelless words, as we do in the present study, is critical in order to make generalizations about what makes them harder (or not) to comprehend above and beyond what a native-speaker experience allows.

1.5. Current study

For the current study, we designed two experiments to investigate the perception of vowelless words in Tashlhiyt by native and non-native listeners. As mentioned above, vowelless words are frequent in Tashlhiyt and vary in length, from two segments (e.g. *fk* 'give') to multisyllabic words containing long strings of consonants (e.g., *tsskffstt* 'you dried it (F)', Ridouane, 2014). In the current study, we focus only on tri-segmental vowelless words (CCC structure) as a test case. The present study had four experimental manipulations. We outline each of these manipulations in turn next, as well as how the perception of vowelless words can shed light on fundamental issues in phonological typology.

First, we varied the listeners' language background: native Tashlhiyt speakers and speakers who are Tashlhiyt-naive (here, American English speakers). English does permit relatively complex syllable structures (e.g., *strengths*, *sixths*) and syllabic consonants (e.g., *bottle*). However, only sonorants in unstressed syllables can be nuclei (though, alternating with a schwa pronunciation, Roach et al., 1992) and no monosyllabic vowelless words are allowed.

Second, we used two different perceptual paradigms: a paired discrimination (4IAX) task and a nonword acceptability ratings task. In a paired discrimination task (Fowler, 1984), a listener hears two pairs of items - one contains the same word repeated twice, the other contains two different words. In our task, the different-pair items varied in the middle segment (e.g., *tuf* vs. *tlf* or *ƀbr* vs. *ƀdr*). The discrimination task gauges low-level, auditory processing of vowelless words since listeners rely on acoustic similarity between pairs. If the cross-linguistic rarity of vowelless words stems from an acoustic processing bias, we predict that discrimination will be lower for vowelless words than for words with vowels. Meanwhile, the wordlikeness task requires listeners to compare the phonetic and phonological properties of stimuli to the characteristics of words in their memory. It has been shown that nonwords with greater lexical

support are rated as more wordlike than nonwords with less (Munson, 2001; Frisch et al., 2001). We predict that Tashlhiyt listeners will show greater perceptual sensitivity and higher wordlikeness ratings of vowelless words than English speakers.

Thirdly, we examine variations in the sonority profile of the first and second segments in the tri-segmental items. We focus on this sonority profile since that is a meaningful classification for Tashlhiyt (Lahrouchi, 2010). Also, word-onset sonority profiles are argued to be more constrained than coda profiles cross-linguistically (Pouplier & Beňuš, 2011), and listeners are more sensitive to sequential probabilities in onset position (van der Lugt, 2001). Tashlhiyt test words contained one of four sonority types (where V = vowel, C = any consonant, O = obstruent, S = sonorant): rising sonority with a vowel nucleus (CVC), vowelless with rising sonority (OSC), vowelless with plateau sonority (OOC; NB: SSC are not frequent forms in Tashlhiyt, Lahrouchi, 2010), and vowelless with falling sonority (SOC). Each of these four sonority profile types are illustrated with examples in Figure 1. In each example, the segments are ranked based on the sonority hierarchy (Parker, 2002) to illustrate the differences across word types. For rising sonority words (both CVC and OSC), sonority increases from the first to the second segment, crossing at least 2 ranks. In plateau sonority words, there is not a large sonority difference from C1 to C2. In falling sonority words, there is a large difference in sonority, but it decreases from C1 to C2. For the ratings task, we constructed Tashlhiyt-like tri-segmental nonwords varying in their structure. By investigating vowelless words that vary in their sonority profile, we can tease apart different theoretical accounts about their perception. On the one hand, as outlined above, it has been argued that the sonority hierarchy is an active mechanism in shaping the perception of sound sequences across languages (Berent et al., 2009). A prediction from such an account is that perceptual sensitivity within vowelless words decreases with decreasing sonority (rising > plateau > falling). On the other hand, sonority *difference* has been argued to be more relevant for perception (Ohala & Kawasaki-Fukumori, 1997). That prediction is that perceptual sensitivity will be higher when words contain greater sonority differences (e.g., rising & falling > plateau). Moreover, differences across tasks can speak to low-level auditory processes (discrimination) vs. higher-level phonological mechanisms (ratings) in the perception of vowelless words.

/sak/ "pass through" CVC rising sonority				/krf/ "tie" OSC rising sonority				/dbɤ/ 'to tan' OOC plateau sonority				/nkr/ 'wake' SOC falling sonority			
V - 8		a		V - 8				V - 8				V - 8			
G - 7				G - 7				G - 7				G - 7			
L - 6				L - 6		r		L - 6				L - 6			r
N - 5				N - 5				N - 5				N - 5	n		
VDF - 4				VDF - 4				VDF - 4			ɸ	VDF - 4			
VLF - 3	s			VLF - 3			f	VLF - 3				VLF - 3			
VDS - 2				VDS - 2				VDS - 2	d		b	VDS - 2			
VLS - 1			k	VLS - 1	k			VLS - 1				VLS - 1			k

Figure 1. Examples of Tashlhiyt words varying in their sonority profiles.

For our fourth manipulation, we elicited all the items for the perceptual experiments from a native Tashlhiyt speaker in both Clear and Casual speaking styles. The style manipulation allows us to explore the relationship between phonetic enhancement and phonological typology. Our predictions for how clear speech variation will influence the perception of vowelless words apply to the discrimination task most readily. One prediction is that vowelless words will receive less of a clear speech boost in perceptual sensitivity than words that contain vowels. Since vowels are the loudest, most sonorous segment, clear speech enhancement can make it even easier for listeners to discriminate words that have different vowels. On the other hand, if the phonetic implementation of vowelless words includes the articulation of salient cues to the internal phonological structure, we predict that the transitional cues that serve to carry information about the segment identities will be enhanced in clear speech. How might clear speech influence nonword acceptability ratings? To our knowledge, no prior study has examined the effect of speaking style on word-likeness judgements. The ratings task can reveal whether speaking style influences listeners' acceptance of different types of nonwords.

2. Experiment 1: 41AX (paired) discrimination

In Experiment 1, listeners completed paired discrimination of tri-segmental word minimal pairs where the middle segment varies.

2.1. Methods

2.1.1. Target words

We compiled a list of CCC and CVC items in Tashlhiyt and identified word pairs contrasting only in the middle segment. There were 23 minimal pairs of two vowelless words contrasting in the middle consonant (CCC vs. CCC), 18 minimal pairs where one was a vowelless word and one word contained a vowel (CVC vs. CCC), and 3 minimal pairs where there was a vowel contrast (CVC vs. CVC) (we additionally identified two additional words that contrasted in vowel quality containing a complex onset: e.g., *lfal* vs. *lfil*). Each minimal pair type could further be classified based on their C1-C2 sonority profile. For CCC vs. CCC, 9 pairs contained obstruents in both C1 and C2 position (plateau vs. plateau pairs), 8 pairs consisted of one word with C1 and C2 obstruents and one word with an obstruent as C1 and a sonorant as C2 (plateau vs. rising pairs), 5 pairs where both contained a sonorant in C1 and an obstruent in C2 (falling vs. falling pairs), and just 1 consisted of one of the pairs having a flat sonority profile containing a sonorant in C1 and C2. For CVC vs. CCC, 11 pairs contained a vowelless word with obstruents in C1 and C2 position (rising CVC vs. plateau sonority), 4 pairs contained a vowelless word with an obstruent as C1 and a sonorant as C2 (rising CVC vs. rising sonority), and 3 contained words with a falling sonority (rising CVC vs. falling sonority).

For the discrimination task, we aimed to select 5 pairs from each of these sonority comparisons as critical trial types (for CVC vs. rising CCC and CVC vs. falling CCC, there were only 4 and 3 pairs, respectively). In addition, 5 sets of non-minimal pair words, containing vowels, were also selected. Selection of the pairs for the discrimination task generated a list of 62 unique lexical items in Tashlhiyt.

In addition to the items selected for the discrimination task, 14 picture-able nouns were selected (*ajjis* 'horse', *tafukt* 'sun', *afullus* 'rooster', *az'alim* 'onion', *ajdi* 'dog', *izi* 'fly', *ajjur* 'moon', *aw^rum* 'bread', *anz^rar* 'rain', *idukan* 'slippers', *tafunast* 'cow', *ilm* 'skin', *tafruxt* 'girl',

afruχ ‘boy’). These words were used in a pre-test language comprehension assessment for the native Tashlhiyt listeners (see section 2.1.6).

2.1.2. Speech styles

All selected words were produced in a randomized order by a native speaker of Tashlhiyt (one of the authors, ML) in two speaking styles. The recording took place in a sound-attenuated booth using an AT 8010 Audio-technica microphone and USB audio mixer (M-Audio Fast Track), digitized at a 44.1 kHz sampling rate. To elicit Clear Speech, the speaker was given instructions similar to those used to elicit clear speech in prior work (e.g., Bradlow, 2002 and Zellou et al., 2022): “In this condition, speak the words clearly to someone who is having a hard time understanding you.” The speaker produced each word in two different frame sentences: *ini* ___ *jat tklit* ‘say ___ once’, *inna* ___ *baħira* ‘he said ___ a lot’.

Following the Clear Speech style elicitation, the speaker produced the words in a fast, casual speaking style with the following instructions also modeled after those used in prior work (e.g., Bradlow, 2002): “now, speak the list as if you are talking to a friend or family member you have known for a long time who has no trouble understanding you, and speak quickly”. The speaker also produced the words in each of the two frame sentences.

2.1.3. Acoustic assessment of the stimuli

Since prior work has shown that clear speech contains longer word durations (Krause & Braidă, 2004), we measured the durations of target words across speaking styles. Table 1 shows the mean and standard deviation of word durations for CVC and CCC items across styles. A two-tailed, unpaired t-test revealed that words are longer in Clear speech than in Casual speech ($t(227) = 19.7, p < 0.001$). We also asked whether vowelless words are overall shorter in duration since they do not contain vowels. A t-test indicated that CCC and CVC items are not different lengths ($t(227) = 0.3, p = 0.76$).

Table 1. Word durations in milliseconds (and standard deviations) for CVC and CCC items produced in clear and casual speech.

	Clear	Casual
CVC	437 (89)	245 (57)
CCC	428 (78)	257 (47)

Ridouane & Fougeron (2011) analyzed the acoustic properties of vowelless words in Tashlhiyt, focusing mainly on sequences of obstruents, and report that a vocoid-like element is often produced between consonants, ranging in quality from a release burst to a full epenthetic vowel. We coded each vowelless word production for whether there was any inter-consonantal release or vocoid-like element either between C1 and C2 or between C2 and C3 (if there were two, we measured the duration of the longer element). We additionally coded whether that element was a full schwa vowel, characterized by higher-level formant structure. We also measured the duration of any vocoid element as the temporal interval between consonants. Table 2 provides the percentage of vowelless words that contain a vocoid-like element, by sonority profile and

across speaking styles. The percentages of those elements that are full schwas across conditions, are also provided, as well as the duration of the element when present.

Table 2. Percentage of tokens with vocoid-like elements, percentage of tokens produced with a full schwa vowel, and durations in milliseconds (and standard deviations) of the element when present, in vowelless words across speaking style.

	% tokens with an inter-consonantal release or vocoid-like element		% tokens with full vowel		Intrusive element duration (when present)	
	Clear	Casual	Clear	Casual	Clear	Casual
Falling	50%	50%	33%	33%	60 (30)	46 (22)
Plateau	75%	75%	44%	34%	45 (26)	31 (15)
Rising	50%	25%	50%	25%	81 (12)	30 (7)

We ran two chi-square tests to examine whether the rates of either any vocoid or the full schwa differs across speaking styles. The tests revealed that the proportion of vowelless tokens with a vocoid-like element is *not* different across clear and casual productions ($X^2(1, N = 132) = 1.1, p = 0.6$) and also that the proportion vowelless words with full vowels produced on vowelless words is *not* different across clear and casual productions ($X^2(1, N = 132) = 0.71, p = 0.7$). However, a two-tailed t-test demonstrated that durations of the vocoid-like element, when present, is longer in clear speech productions ($t(71) = 2.1, p < 0.05$). Thus, vocoid-like elements, or full epenthetic schwas, are often produced in vowelless words (consistent with Ridouane & Fougeron, 2011). While vocoid elements occur at the same rate in clear and casual speech, they are more likely to be longer in clear speech when they are present.

2.1.4. Stimuli preparation

All items were segmented and excised from their frame sentences and amplitude normalized to 65 dB. Tokens for each trial were concatenated into a single sound file with a within-pair inter-stimulus interval of 300 ms and a pair-medial ISI of 500 ms. Typically, discrimination paradigms present speech embedded in noise in order to increase the difficulty of the task. Therefore, all stimuli were mixed with white noise (which has been shown to mask consonants more uniformly than other types of noise, Phatak & Allen, 2007) at a signal-to-noise ratio (SNR) of 0 dB (Miller & Nicely, 1955).

2.1.5. Listener participants

36 native Tashlhiyt speakers (11 female, 0 non-binary/other, 23 male (2 did not report their gender); mean age = 40.2 years old) and 36 English speakers (21 female, 2 non-binary/gender-fluid, 13 male; mean age = 20.2 years old) completed the online experiment. All participants completed informed consent before participating. None of the listeners reported having a hearing or language impairment.

The Tashlhiyt participants were recruited through email flyers. All the Tashlhiyt participants reported that Tashlhiyt was their first language and that both parents speak Tashlhiyt. Participants also reported that they spoke French (all instructions were provided in French), and Arabic (Moroccan and Standard). Some reported to also speak additional languages (English, $n = 22$; Spanish, $n = 1$; Italian, $n = 1$; German, $n = 1$; Turkish, $n = 1$). Participants reported growing up in cities such as Agadir ($n = 16$), Marrakech ($n = 1$), Essaouira ($n = 2$), Tiznit ($n = 4$), or other towns and villages in Southern Morocco. The English-speaking participants were recruited from the [redacted for anonymity] subjects' pool. All of the participants reported being native speakers of American English. Four participants reported that they speak a language other than English in the home (Swahili, $n = 1$; Estonian, $n = 1$; Vietnamese, $n = 1$; Punjabi, $n = 1$). We asked the English-speaking participants if they spoke or had studied Tashlhiyt or any of the languages of North Africa; none reported that this was the case.

2.1.6. Paired discrimination (41AX) task and procedure

The experiment was conducted online using Qualtrics. Participants were instructed to complete the experiment in a quiet room without distractions or noise, to silence their phones, and to wear headphones. As most computer-literate Tashlhiyt speakers are fluent in French, all instructions in the experiment were presented in French to the Tashlhiyt participants. All instructions were provided in English to the English-speaking participants.

For the Tashlhiyt participants, the experiment began with a pre-test word identification task in order to assess that they were speakers of Tashlhiyt. On each pre-test trial, listeners heard one of 14 lexical comprehension nouns and were presented with two black and white images on the screen - one depicted the target word. Participants selected the image corresponding to the word they heard. Participants heard each of the nouns once, randomly presented, completing 14 of these trials. All of the Tashlhiyt participants correctly identified all the nouns in these comprehension trials.

For the English-speaking participants, the study began with a pre-test of their audio: participants heard one sentence presented auditorily (“She asked about the host”) and were asked to identify the sentence from three multiple choice options, each containing a phonologically close target word (host, toast, coast). Participants were also presented with six trials of a tone-identification task (Woods et al., 2017). All participants passed these tasks.

After the pre-test procedure, all participants completed a paired (41AX; binary forced-choice) discrimination task. In each trial, two pairs of words are played. One pair contains different words (e.g., *tuf* vs. *tlf*), and the other pair contains different productions of the same word (e.g., *tlf* vs. *tlf*). Participants were asked which pair contains different words (for the Tashlhiyt participants: “Quelle paire est constituée de deux mots différents?”; for the English-speaking participants: “Select which pair of words sound most different”), and identify either the first pair (“la première paire” | “First Pair”) or the second pair (“la deuxième paire” | “Second Pair”) via a mouse click. The Tashlhiyt listeners were told they would hear words in Tashlhiyt. The English-speaking participants were told that they would be hearing utterances produced in a language they have never heard before and that it is fine that they don't understand the words. The stimuli in each trial were presented once only, with no possibility to repeat the sound file in a trial. Before starting the task, participants were given written instructions for this task and they

completed two practice trials, one where the first pair contained different words and one where the second pair contained different words, with feedback on the correct responses for those trials.

Table 3 lists the different trial types used in the paired discrimination task and provides examples of each (a full list of all of the word pairs is provided in the Appendix). Across trials, speech style varied: each of the 37 trial types were randomly presented to the listeners in both clear and casual speech conditions (74 total discrimination trials). The ordering of words within different pairs, identity of same pair words, and ordering of different and same pairs varied equally across trial types within-subject and were counterbalanced across four experiment lists. The order of trials was randomized for each participant. Following the discrimination task, participants also completed a nonword acceptability task (described in Section 3.1). In total, the entire study took approximately 30 minutes to complete.

Table 3. Examples of the different paired discrimination trial types (the full list of words used in trials, with their glosses is provided in the Appendix).

minimal pair type	Sonority Comparison (n of types)	Different Pair example	Same Pair example
non-MPs	non-minimal pairs (5)	<i>r⁶u^h</i> vs. <i>sir</i>	<i>sir</i> vs. <i>sir</i>
CVC vs. CVC	risingCVC vs. risingCVC (5)	<i>fan</i> vs. <i>fin</i>	<i>fin</i> vs. <i>fin</i>
CCC vs. CVC	risingCVC vs. rising (4)	<i>tuf</i> vs. <i>tlf</i>	<i>tlf</i> vs. <i>tlf</i>
	risingCVC vs. plateau (5)	<i>fat</i> vs. <i>fkt</i>	<i>fkt</i> vs. <i>fkt</i>
	risingCVC vs. falling (3)	<i>r⁶u^h</i> vs. <i>rb^h</i>	<i>rb^h</i> vs. <i>rb^h</i>
CCC vs. CCC	rising vs. plateau (5)	<i>zlm</i> vs. <i>zdm</i>	<i>zdm</i> vs. <i>zdm</i>
	falling vs. falling (5)	<i>ns⁶h</i> vs. <i>n3^h</i>	<i>n3^h</i> vs. <i>n3^h</i>
	plateau vs. plateau (5)	<i>ɸbr</i> vs. <i>ɸdr</i>	<i>ɸdr</i> vs. <i>ɸdr</i>

2.2. Results

The data from the 72 listener responses were coded for correctly (=1) or incorrectly (=0) selecting the pair that contained different words in each trial and modeled using a mixed effects logistic regression with the *glmer()* function in the *lmer* package (Bates et al., 2015) in R (R core team). The model included three fixed effects (all treatment coded): Language Background (Tashlhiyt, English [reference level]), Speech Style (Clear, Casual [reference level]), and Minimal Pair Type (Non-minimal pair [reference level], CVC vs. CVC, CVC vs. CCC, CCC vs. CCC). The random effects included random intercepts for each participant, word pair type, and word pair ordering, as well as by-participant random slopes for Speech Style (including by-participant random slopes for Minimal Pair Type resulted in a singularity error indicating over fitting).

The addition of 2- and 3-way interactions between predictors was evaluated via model comparison using the *anova()* function based on the Akaike Information Criterion (AIC), measuring model goodness of fit while penalizing over-parameterization. The model that results in the lowest AIC is best supported by the data (Akaike, 1974). The model with two-way interactions between Minimal Pair Type and Style as well as Language Background and Style had the lowest AIC (AIC = 4254.1, model ANOVA Chisq = 7.3, df = 1, $p < .01$) and was thus

retained. Inclusion of the three-way interaction did not improve model fit. Prior work has used a predictor-releveling procedure for interpreting *lmer* models that contain a factor with more than two levels. In this method, the summary statistics for a model run on the same data set is output with different reference levels for the factor of comparison (Clopper, 2013; Tremblay et al., 2017). We followed this procedure: the model was run with the reference level for Minimal Pair Type re-leveled to produce estimates for multiple pairwise contrasts. (Output 1 reference level = Non-minimal pairs, Output 2 reference level = CVC vs. CVC; Output 3 reference level = CCC vs. CCC) The summary statistics from the model (and relevelled outputs) is provided in a table in the Appendix.

Figure 2 shows the mean discrimination for each Minimal Pair Type by Speaking Style, across listener groups. The model revealed a main effect of Minimal Pair Type. Discrimination performance is highest in trials where the different pair contained a non-minimal pair (91% accuracy) compared to trials with minimal pairs, which are not overall different from one another (CCC vs. CCC: 79%; CVC vs. CVC: 77%; CVC vs. CCC: 81%). In the models where minimal pairs were the reference level, there was a main effect of style wherein discrimination performance is higher for items produced in Clear speech than in Casual speech.

There was also an interaction between Minimal Pair Type and Speech Style. As seen in Figure 2, the difference between discrimination performance for Clear and Casual speech styles was largest for the CVC vs. CVC comparison. As seen in Figure 2, this is realized as the lowest discrimination performance for CVC vs. CVC trials in Casual speech, while performance is more similar across minimal pair types in Clear speech.

Language Background of the listener also predicted overall discrimination performance: Tashlhiyt listeners (87%) are better at discrimination of Tashlhiyt words than English-speaking listeners (76%). There was also an interaction between Language Background and Style: the positive coefficient value indicates that Tashlhiyt listeners perform better in Clear speech than the naive listeners.

As mentioned above, the inclusion of the three-way interactions with Language Background and the other predictors did not significantly improve model fit. This indicates that, while English listeners overall perform lower on discrimination of Tashlhiyt words than native listeners, they do not display even lower performance of a function of any of the word or speech style types in the experiment.

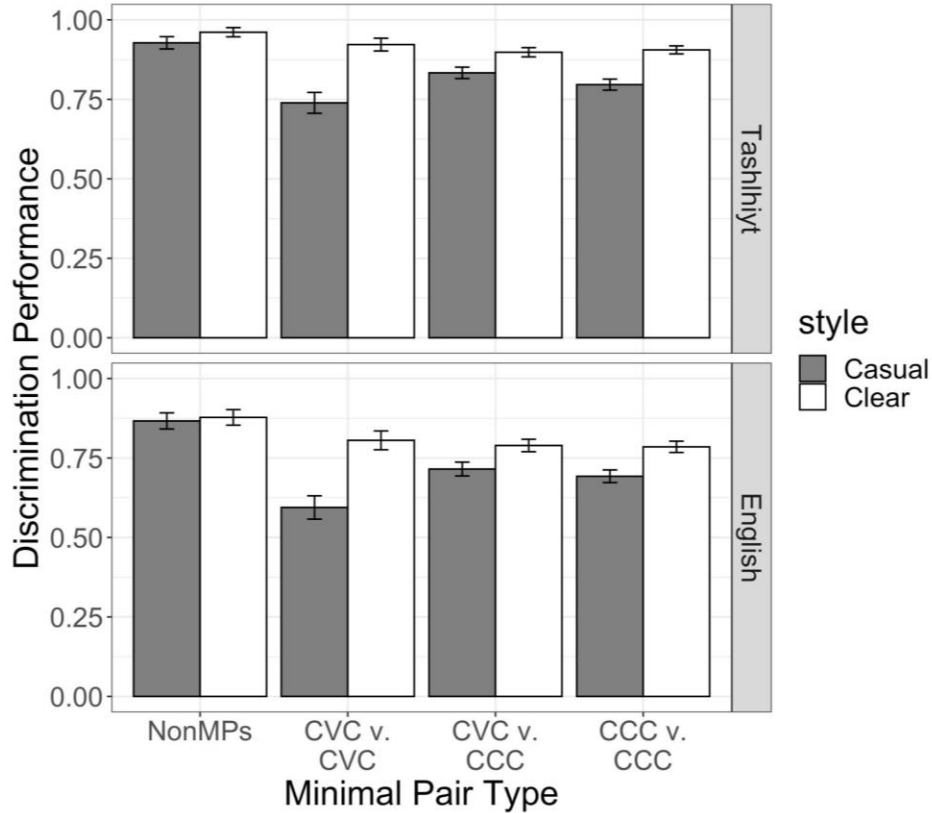


Figure 2. Experiment 1. Performance on paired discrimination trials across Minimal Pair and Speech Style conditions for native Tashlhiyt and English listeners.

We also explore the extent to which sonority patterns *within* vowelless words lead to differences in perceptual sensitivity. To that end, two post-hoc logistic regression models were run separately on the CVC vs. CCC trials subset and the CCC vs. CCC trials subset. Both models contained three fixed effects. The first two effects were the same across the two post-hoc models and coded identically as in the main model: Language Background (Tashlhiyt, English) and Style (Clear, Casual). The third fixed effect, Minimal Pair Type, was coded differently across the two models. In the CVC vs. CCC model, Minimal Pair Type contained 3 levels: CVC vs. CCC-Falling, CVC vs. CCC-Plateau, and CVC vs. CCC-Rising. In the CCC vs. CCC model, the three levels of the predictor Minimal Pair Type were CCC-Falling vs. CCC-Falling, CCC-Plateau vs. CCC-Plateau, and CCC-Rising vs. CCC-Rising. The random effects structure of both models was identical to that for the main model. As with the main model, we performed model comparisons to evaluate the inclusion of each fixed effect and all two- and three-way interactions. For the CVC vs. CCC model, only inclusion of the two-way interaction between Minimal Pair Type and Style was computed as the best fit and thus retained (model ANOVA $\text{Chisq} = 6.8$, $\text{df} = 2$, $p = .03$). For the CCC vs. CCC model, inclusion of interactions between Minimal Pair Type and Style and Language Background and Style resulted in the best fit (model ANOVA $\text{Chisq} = 4.5$, $\text{df} = 1$, $p = .03$).

The summary statistics for both models are provided in the Appendix. Figure 3.A shows the mean proportion of correct discrimination for CVC vs. CCC trials. The model revealed an interaction between Speech Style and Minimal Pair Type: the difference in accuracy between Clear and Casual items is larger for the CVC vs. CCC-Falling trials than for the CVC vs. CCC-

rising trials. This interaction can be seen in Figure 3.A: clear speech boosts the discrimination between CVC vs. CCC-Falling trials to the greatest extent.

Fig. 3.B shows the mean proportion of discriminated trials containing only vowelless words. For the CCC vs. CCC model, there was a main effect of Minimal Pair Type wherein trials containing vowelless words with falling sonority profiles are the hardest to discriminate. There was also a significant interaction between Minimal Pair Type and Speech Style. As seen in Fig 3.B, the size of the increase in discrimination performance from casual to clear speech is larger for trials with plateauing sonority vowelless words, compared to trials containing a plateau and rising sonority vowelless words.

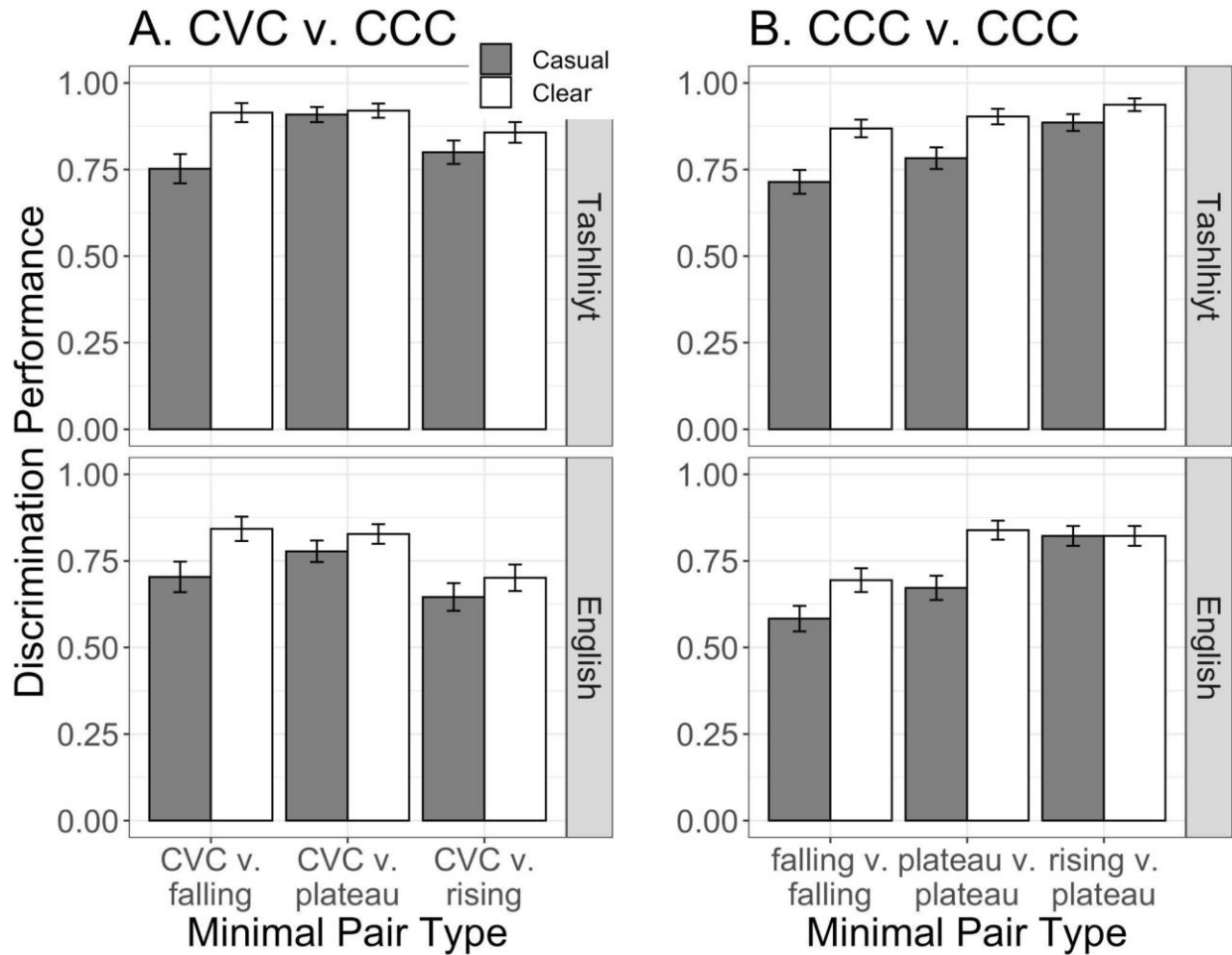


Figure 3. Experiment 1. Performance on paired discrimination trials with vowel and vowelless items (Panel A) and vowelless words only (Panel B) across Sonority Profile and Speech Style conditions for native Tashlhiyt and English listeners.

2.3. Interim discussion

In Experiment 1, native Tashlhiyt and Tashlhiyt-naive (English-speaking) listeners show variation in discrimination performance for different types of minimal pair contrasts. As expected, minimal pairs are harder to discriminate between than non-minimal pairs. However, within minimal pair trials, perceptual sensitivity to vowelless minimal pairs (contrasting in the

center segment) is not different overall than for words containing vowels, as well as vowelless words are not harder to discriminate between than words with vowels.

We also found that vowelless word pairs containing falling sonority profiles (e.g., *nsʰ* vs. *nʒh*) are harder to discriminate than pairs containing plateaued sonority (e.g., *ɤbr* vs. *ɤdr*). This result is surprising considering proposals that consonant sequences containing larger sonority differences will be easier to discriminate than those with smaller sonority differences (Ohala & Kawasaki-Fukumori, 1997). However, the lowest performance for falling sonority contrasts is consistent with the observation that falling sonority clusters are the least common cross-linguistically and prior findings that they are most perceptually dispreferred by listeners from a range of language backgrounds (Berent et al., 2009).

Notably, while non-native listeners performed overall lower than native Tashlhiyt listeners, the perceptual discrimination patterns within each contrast was parallel across listener groups. Thus, the lack of native language experience with Tashlhiyt does not make vowelless words less discriminable relative to other Tashlhiyt words. Relative perceptual sensitivity to vowelless words varying in sonority patterns was also similar across native and non-native listeners, further supporting that cross-language principles influence discriminability of vowelless words.

Additionally, vowelless words produced in Clear speech are better discriminated than when those words are produced in a more reduced speaking style. The biggest clear speech perceptual boost is for minimal pairs contrasting in vowels. This is not surprising since vowels are more informative than consonants. However, it is unexpected that trials containing CVC pairs in reduced speech had the lowest discrimination performance (hence, clear speech is actually enhancing the discriminability of voweled words to be equal to other minimal pair trial types). In the general discussion, we explore possible explanations for this finding. We also observe that, within vowelless words, clear speech boosts discriminability of the harder sonority contrasts to a greater extent. In particular, falling and plateau sonority vowelless words are perceptually enhanced. Moreover, enhancement is observed for both listener groups. This adds to growing empirical work about how listener-oriented speaking styles can impact the perceptibility of a range of cross-linguistic phonological patterns (Kang & Guion, 2008; Smiljanić & Bradlow, 2005).

3. Experiment 2: Nonword Acceptability Ratings

Experiment 2 is a nonword acceptability judgment task with three-segment nonwords where the sonority value of the second segment varies.

3.1. Methods

3.1.1. Nonwords

In Tashlhiyt, any consonant can occupy C2 in a CCC word (with the exception of glides, for which only one real word was identified, *rwɪ*). Thus, the non-words were constructed to be possible Tashlhiyt nonce words containing the most common C1s (*f, n, r*) and C3s (*l, r, n, m*) found in Tashlhiyt, but varying in the sonority value of C2. The nonwords created for Experiment 2 are provided in Table 4.

Table 4: Nonwords for Word Acceptability Ratings Task.

Nucleus	Nonword frame: /f_r/	Nonword frame: /n_l/	Nonword frame: /r_n,m/
vowel	<i>fur</i>	<i>nul</i>	<i>run</i>
liquid	<i>flr</i>	<i>nrl</i>	<i>rln</i>
nasal	<i>fnr</i>	<i>nml</i>	<i>rmn</i>
voiced fricative	<i>fzr</i>	<i>nɸl</i>	<i>rɸm</i>
voiceless fricative	<i>fçr</i>	<i>nçl</i>	<i>rçn</i>
voiced stop	<i>fdr</i>	<i>ndl</i>	<i>rdn</i>
voiceless stop	<i>fqr</i>	<i>nkl</i>	<i>rtm</i>

3.1.2. Speech styles and stimuli preparation

The same speaker from Experiment 1 also produced the nonword items. A trained linguist, the speaker read a list of the nonwords transcribed in IPA. The speaker produced all the nonwords in clear and casual speaking styles following the same instructions and procedures from Experiment 1. The nonwords were extracted from their frame sentences and amplitude normalized to 65 dB. The first repetition of each nonword in each speaking style was selected as the stimuli.

3.1.3. Participants and procedure

The same 36 Tashlhiyt and 36 English-speaking participants from Experiment 1 also completed Experiment 2. Each trial consisted of the auditory presentation of a nonword, once only with no option to repeat. Listeners were instructed to rate how likely the word they heard could become a word in their language in the future (either Tashlhiyt for the Tashlhiyt participants, or English for the English participants: “Évaluez chaque mot et marquer sur l'échelle de 0 à 100 dans quelle mesure il peut être un mot possible du tachelhit dans l'avenir proche” | “Rate how likely you think this word could become a new word in English in the future”) (Daland et al., 2011). Participants marked their rating on a sliding scale from 0 (“impossible” | “not at all likely”) to 100 (“tout à fait possible” | “very likely”) (the default position of the marker was reset to the midpoint (50) at the start of each trial).

Two experimental lists were constructed. In the first list, half of the nonwords were presented in the clear speaking style, while the other half of the nonwords were presented in the casual style. Style assignment was counterbalanced across words in the second list. Thus, participants heard each nonword only once (21 trials total), presented in either the clear or casual styles. Trial order was randomized for each participant.

3.2. Results

We modeled nonword acceptability ratings as a continuous dependent variable (0-100; centered and scaled prior to model fitting) with a linear mixed effects model using the *lme4* R package. The model included three fixed effects: Language Background (Tashlhiyt, English [reference

level]; treatment coded), Speech Style (Clear, Casual [reference level]; treatment coded), and Sonority value of the middle segment. We adopt the sonority scale of Parker (2002) where integers are assigned to sounds based on their sonority ranking (vowel = 8, glide = 7, liquid = 6, nasal = 5, voiced fricative = 4, voiceless fricative = 3, voiced stop = 2, voiceless stop = 1). Sonority value of the middle segment was a continuous predictor, centered and scaled prior to model fitting. The random effects included by-participant and by-item random intercepts, as well as by-participant random slopes for Sonority Value and Speech Style (including by-participant random slopes for the interaction between Sonority Value and Speech Style resulted in a singularity error indicating over fitting).

Just as we did for Experiment 1, the addition of 2- and 3-way interactions between predictors was evaluated via model comparisons. The model with all two- and three-way interactions had the lowest AIC (model ANOVA Chisq = 20.4, $df = 1$, $p < 0.01$) and was thus retained.

A Table with the model output is provided in the Appendix. The model revealed a main effect of Language Background wherein Tashlhiyt listeners provide higher acceptability ratings of the nonwords than English listeners (Tashlhiyt mean = 54.8, English mean = 36.6). There was also an effect of Speech Style: nonwords produced in Clear speech have higher ratings (Clear = 49.8, Casual = 41.6). There was also an interaction between Sonority value of the middle segment and Speech Style. In general, nonwords with a higher sonority middle segment are rated as more word-like if they are produced in Clear speech.

However, there was also a three-way interaction between listener Language Background, Sonority Value, and Speech Style. This interaction is depicted in Figure 4, which displays the mean nonword acceptability ratings by middle segment and Speaking Style, across Tashlhiyt and English-speaking listeners. As seen, English listeners are more likely to provide higher nonword acceptability ratings as the middle segment increases in Sonority - but the increase is only for items produced in Clear speech; acceptability ratings are low for phonetically reduced items across the board for the Tashlhiyt-naive listeners. Meanwhile, Tashlhiyt listeners display the reverse sonority-based acceptability pattern: nonwords containing middle segments that decrease in sonority are more likely to receive higher word-like judgements, especially when they are produced in Clear speech. Figure 5 provides the mean nonword acceptability ratings across speech styles and listener groups, but classifying the middle segment as either a vowel, sonorant, or obstruent. The interaction is more visually apparent in Figure 5: while English listeners give higher ratings to vowel and sonorant-nuclei items produced in Clear speech, Tashlhiyt listeners show preference for obstruent-centered words when they are produced in Clear speech.

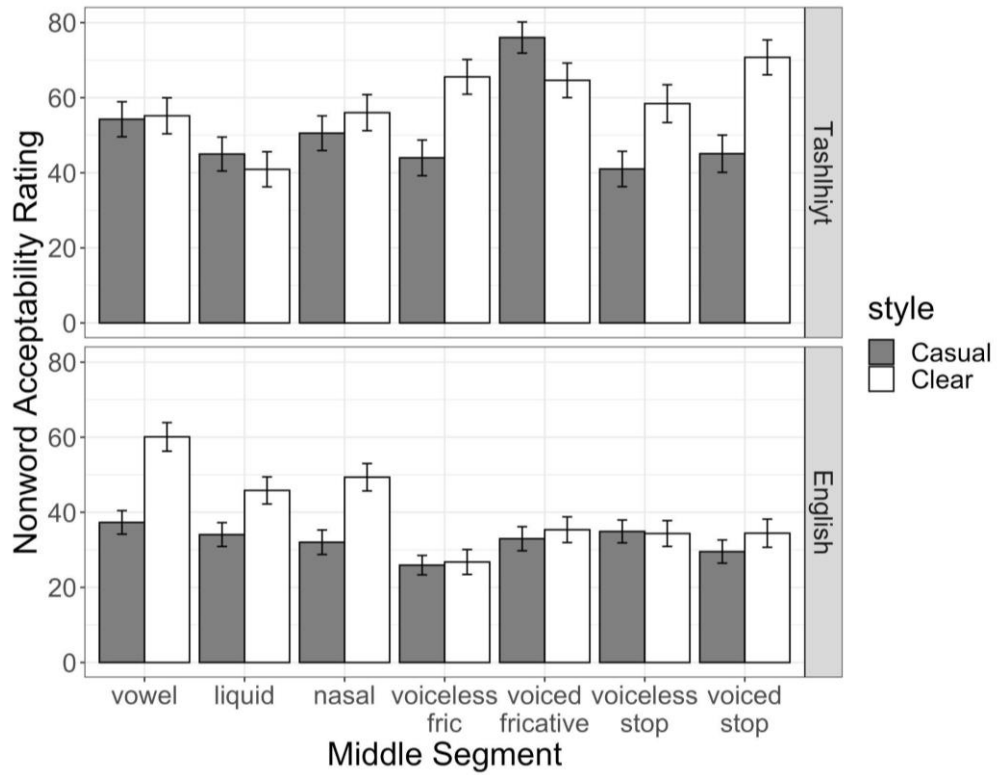


Figure 4: Nonword acceptability ratings by Middle Segment Type, by Speech Style for Tashlihiyt and English listeners.

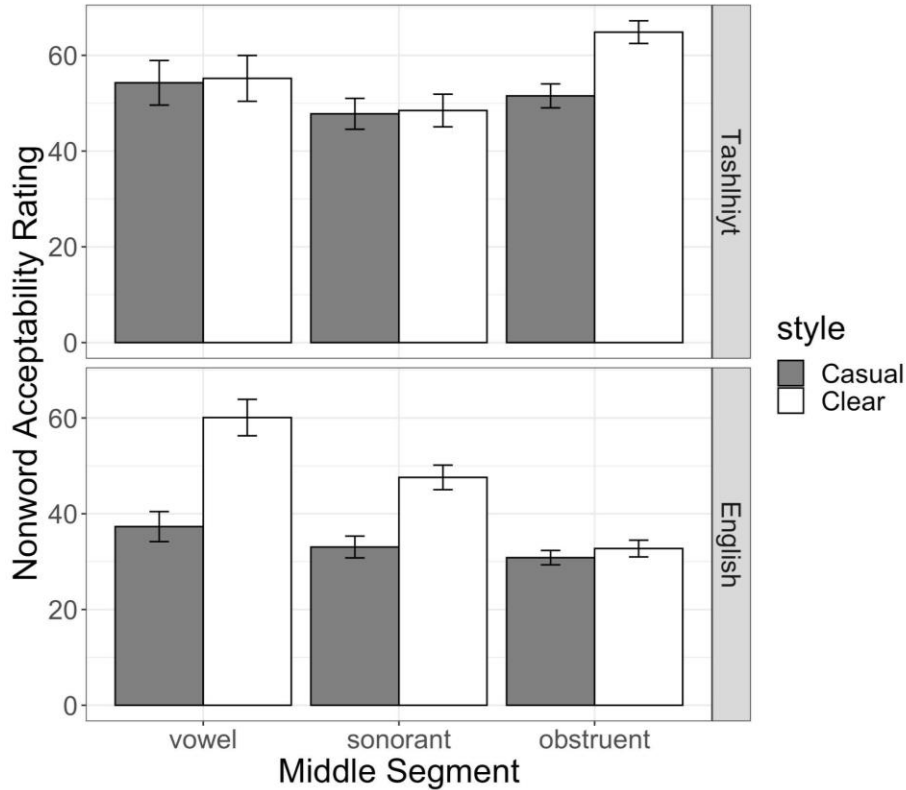


Figure 5: Nonword acceptability ratings where the middle segment is classified as a vowel, sonorant, or obstruent, by Speech Style for Tashlhiyt and English listeners.

4. General discussion

Discrimination responses reveal that both native and non-native listeners are generally able to identify Tashlhiyt vowelless words contrasting in the middle consonant. There is variation in perceptual sensitivity within vowelless words, however, based on the sonority patterns of the initial consonants: words with falling sonority sequences (e.g., *nkr* vs. *ngr*) are harder to discriminate than plateauing (obstruent-obstruent) sonority words (e.g., *ɛbr* vs. *ɛdr*). However, Tashlhiyt- and English-speaking listeners display distinct rating patterns of Tashlhiyt-like nonwords: native listeners are more likely to prefer obstruent-centered vowelless words, while non-native listeners prefer words with vowels. Across experiments, perceptual patterns are influenced by speaking style: in some cases, clear speech increases discriminability of and nonword acceptability ratings for vowelless words. In this discussion, we examine our results and consider their implications for phonological theory and phonotactic typology.

4.1. Perception of vowelless words in Tashlhiyt

Vowelless words are rare in languages of the world. However, we find that they are not more difficult to discriminate in Tashlhiyt. This is surprising given that vowels provide robust transitional coarticulatory cues about the identity of adjacent sounds (Lieberman et al., 1954) and that consonant clusters are harder to perceive than single consonant onsets (Foss & Gernsbacher, 1983; Cutler et al., 1987). On the other hand, since Tashlhiyt has a large consonant inventory,

there is pressure to enhance acoustic cues to segment identity in vowelless words, even in casual speech where consonant cues are more likely to be obscured. Indeed, articulatory work shows that consonant sequences are non-overlapping in Tashlhiyt and produced with transitional vocoid elements that could provide acoustic cues to the segmental composition of vowelless words (Ridouane & Fougeron, 2011). Moreover, the acoustic analysis of our stimuli revealed that vocoid-like elements, and even full schwas, are frequent in the vowelless items and occur at similar rates across casual and clear speech. Thus, it appears that speakers do provide cues about the internal structure of vowelless words in Tashlhiyt even when producing reduced speech. This parallels findings that word-initial stop+stop sequences in Tsou are produced with audible release bursts frequently and at similar rates across slow, normal, and fast speaking rates (Wright, 1996). Our present findings are consistent with a *cue enhancement* view that vowelless words in Tashlhiyt are produced in such a way to provide salient acoustic cues that help listeners recover segmental differences within them.

Furthermore, within vowelless words in Tashlhiyt, we found variation in perceptual patterns based on the sonority profile of consonant sequences. We find that when there is matching word-onset sonority across vowelless pairs (here, falling vs. falling or plateau vs. plateau), discrimination is harder than when sonority varies (rising vs. plateau). In other words, greater acoustic similarity across vowelless words makes discrimination more difficult (Ohala & Kawasaki-Fukumori, 1997). Discrimination performance for sonority-matching pairs is even lower in casual speech. Thus, reduced speech makes it even harder to perceive acoustically similar vowelless words. However, in clear speech, matching plateau sonority items are not harder to discriminate than when vowelless pairs that contain different sonority profiles. There is also a boost for the matching falling sonority items. Thus, clear speech does provide a boost to the harder-to-discriminate vowelless contrasts. Our acoustic analysis suggests that the *length* of the transitional elements is enhanced in clear speech. This adds to cross-linguistic evidence that clear speech provides more auditorily robust phonetic variants of words (e.g., Smiljanić & Bradow, 2005), extending this to vowelless words in Tashlhiyt.

4.2. Perception of words with vowels in Tashlhiyt

We also observed that Tashlhiyt words containing vowels receive the lowest discrimination performance in reduced speech. Why are voweled words more difficult to perceive than vowelless words? One possibility is that strong consonant-to-vowel coarticulatory influences in Tashlhiyt lead to substantial variation across vowels in CVC words in reduced speech such that discriminating between different vowels is difficult. There is some work suggesting that consonant-to-vowel coarticulation leads to a large amount of phonetic variation in vowels for languages with a high consonant-to-vowel ratio, like Tashlhiyt. For example, it has been shown that there is substantial acoustic variation in vowels based on consonantal context in Arabic (also a language with a 3-vowel vs. many consonant inventory) (e.g., Embarki et al., 2007; Bouferroum & Boudraa, 2015). High vowel variability due to consonant-to-vowel coarticulation is reported for other languages with larger consonant-to-vowel ratios (e.g., Salish in Bessell (1998), Yanyuwa and Yindjibarndi in Tabain & Butcher (1999)).

Notably, the difference in perceptual performance across voweled and vowelless words in Tashlhiyt indicates greater consonant contrast preservation than vowel phoneme preservation. This can be explained by the gestural coordination patterns in the language: consonant-to-consonant coarticulation is minimal, but there is extensive consonant-to-vowel coarticulation. In

effect, then, vowels in Tashlhiyt are not as acoustically informative as consonants. Note, however, that the discrimination conditions containing vowelless minimal pairs (CVC vs. CCC) are equivalent in performance with vowelless trials. Within those trial types, words with similar acoustic profiles (CVC vs. CCC-rising) are more difficult to discriminate in both reduced and clear speech, and those with the greatest acoustic difference in their sonority profile (CVC vs. CCC-plateau) have the highest performance across speech styles. Accordingly, Tashlhiyt vowels contrasting with a (obstruent) consonant are easier to discriminate, as opposed to vowels contrasting with sonorants and other vowels, which are hard to discriminate. Work exploring the production and perception of vowel variation in languages with high consonant-to-vowel ratios is a ripe avenue for future research.

4.3. Comparison of native and naive listeners

We also compared native and naive groups of listeners. Tashlhiyt listeners display overall higher discrimination performance than English listeners; however, the relative patterns across word types are basically the same for the two listeners groups. Thus, discrimination of vowelless words in Tashlhiyt, in particular, does not appear to be modulated by experience with the language. In contrast, there are robust differences in Tashlhiyt nonword acceptability ratings across listener groups. Tashlhiyt listeners are more accepting of vowelless nonwords with an obstruent center, while English listeners prefer nonwords with vowel nuclei. These cross-language differences can be related to listener native language-experience: vowelless words with obstruent centers are the most common structure in Tashlhiyt CCC words (around 32% of tri-segmental verbs in Lahrouchi, 2010), and English does not allow vowelless words (a preference for sonorant-centered over obstruent-centered nonwords by English listeners perhaps also reflects the allowance of syllabic sonorants in English, even though this is permitted only in unstressed syllables).

We also found that the clear speech benefit in the discrimination task is larger for Tashlhiyt listeners than for English-speaking participants. This aligns with prior findings that clear speech enhances intelligibility for native speakers to a greater extent than non-native listeners (Bradlow & Bent, 2002). Also notable is that the largest cross-group differences in the ratings task are for clear speech. This is the first study to our knowledge comparing wordlikeness ratings across items produced in different speaking styles. The clear speech forms are easier to parse, phonologically, and provide more intelligible words that allow for phonotactic preferences to emerge. In this way, while some paradigms might make clear speech appear to be “native-listener oriented” (cf. Bradlow & Bent, 2002), we find there are some tasks where the perceptual benefit of clear speech is greater for non-native listeners; here, for the rating of vowelless nonwords as being more like a word of English.

More broadly, that we observe overall qualitative differences in the native and naive listeners’ performance across perception tasks can also speak to the phonological processing of vowelless words. Discrimination relies more on low-level, auditory mechanisms and differences within vowelless words are paralleled across the listener groups. Thus, besides the overall greater perceptual sensitivity by native listeners, experience with Tashlhiyt does not qualitatively change a listeners’ ability to discriminate when the word pairs are acoustically distinct enough. However, for the wordlikeness task, performance recruits higher-level phonological processing mechanisms, and native-language-specific phonotactic patterns influence the patterns across listener groups. Thus, the perception of vowelless words has both a cross-linguistic, auditorily-

based processing mechanism and a phonological component that recruits native-language-specific phonotactic patterns.

4.4. Implications for phonological typology

If vowelless words are not categorically difficult to discriminate, why are they so cross-linguistically rare? Articulatory and acquisitional influences are also argued to shape the evolution of consonant sequencing preferences (Ohala & Kawasaki-Fukumori, 1997; Blevins, 2004). For example, children's first words are typically CV(C), despite hearing words that contain more complex structures (de Boysson-Bardies, 1999; McLeod, Doorn, & Reed, 2001). For articulation, speech production models propose that CV sequences have a unique coordination relationship with one another, which makes that structure particularly stable (Saltzman & Byrd, 2000; Tilsen, 2016). Also, adjacent consonants with very different places of articulation, for instance, might be articulatorily difficult because the tongue configurations might be incompatible (Ohala, 1983). Therefore, articulatory biases could explain why vowelless words are not more common cross-linguistically.

Moreover, as mentioned above, Tashlhiyt has a unique phonetic profile - particularly wide, non-overlapping consonant sequences with frequent audible, vocoid-like elements - which make consonant sequences recoverable to listeners. "Noncoarticulating" consonant clusters don't appear to be uncommon, but complex clusters with heavy overlap appear to be more frequent across languages where the articulatory dynamics of consonant sequences has been studied (Hermes et al., 2017). It is possible that the phonetic pre-conditions of non-overlapping consonants with audible transitions permit the maintenance of vowelless words. These speculations open lines of inquiry that can be explored in future work.

Notably, given the phonetic implementation of consonants in Tashlhiyt, plateauing sonority vowelless words (containing obstruent-obstruent sequences) are easier to discriminate than words with falling sonority, even though the latter contain larger sonority differences between segments. This finding is inconsistent with claims that universal sonority constraints actively govern the structure of words (Berent et al., 2009). Rather, as we see in Tashlhiyt, language-specific phonetic implementation of sound sequences influences the synchronic stability of complex phonotactic structures. Future work directly comparing the perception of consonant clusters across languages that are more coarticulating (e.g., English) vs. less coarticulating (like Tashlhiyt) can provide further insight into how cross-linguistic differences in articulatory dynamics might make some sound contrasts more or less stable. Moreover, while we used only naturally-produced tokens in the present study, manipulation of the phonetic properties of the vocoid elements within Tashlhiyt can be a further testing ground for understanding the relationship between phonetic variation and perceptual sensitivity of vowelless words.

5. Conclusion

The present study examined the perception of vowelless words in Tashlhiyt. We found that due to their phonetic implementation of systematic transitional vocoid elements, listeners can discriminate between vowelless words. Our results from a wordlikeness task reveal the influence of native-language experience on the acceptance of Tashlhiyt-like words varying in sonority profiles. Thus, there are both low-level, auditory and higher-level, memory-based processes involved in the perception of vowelless words. This observation contributes to understanding the

relationship between language-specific phonetic/phonological patterns and the perceptual processing of different types of word forms.

Acknowledgements

Competing interests

The authors have no competing interests to declare.

Ethics and consent

Research was performed in accordance with the XX (IRB number 1864401-1).

APPENDIX

Table X.

Word Type	Sonority Comparison	Word 1	Word 2	Word 1	Word 2
non-mps	non-mps	<i>r^hu^h</i> (go home)	<i>sir</i> (go!)	c-v-c	c-v-c
non-mps	non-mps	<i>lu^h</i> (throw)	<i>ran</i> (they want)	c-v-c	c-v-c
non-mps	non-mps	<i>z^hur^h</i> (visit)	<i>sak</i> (pass through)	c-v-c	c-v-c
non-mps	non-mps	<i>sul</i> (stay alive)	<i>mit</i> (what)	c-v-c	c-v-c
non-mps	non-mps	<i>zud</i> (like, as)	<i>li^h</i> (I married)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	<i>fan</i> (they gave)	<i>fin</i> (they suppurated)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	<i>man</i> (which)	<i>mun</i> (accompany someone)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	<i>kar</i> (only)	<i>kir</i> (only)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	<i>l^hdid</i> (new)	<i>l^hdud</i> (ancestors)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	<i>l^hfal</i> (omen)	<i>l^hfil</i> (elephant)	c-v-c	c-v-c
CCC-CVC	risingCVC v rising	<i>fr^h</i> (be happy)	<i>fu^h</i> (to revel in)	o-s-o	o-v-o
CCC-CVC	risingCVC v rising	<i>ʕlf</i> (feed)	<i>ʕif</i> (to get tired of)	o-s-o	o-v-o
CCC-CVC	risingCVC v rising	<i>tlf</i> (get mixed up)	<i>tuf</i> (she's better)	o-s-o	o-v-o
CCC-CVC	risingCVC v rising	<i>slt</i> (leave on the sly)	<i>sut</i> (drink it)	o-s-o	o-v-o
CCC-CVC	risingCVC v plateau	<i>kdr</i> (betray)	<i>kar</i> (only)	o-o-s	o-v-s
CCC-CVC	risingCVC v plateau	<i>kbr</i> (disappear)	<i>kar</i> (only)	o-o-s	o-v-s
CCC-CVC	risingCVC v plateau	<i>fkt</i> (give it)	<i>fat</i> (give-2MS.PL)	o-o-o	o-v-o
CCC-CVC	risingCVC v plateau	<i>fth</i> (to operate)	<i>fu^h</i> (to revel in)	o-o-o	o-v-o
CCC-CVC	risingCVC v plateau	<i>db^h</i> (tan)	<i>da^h</i> (again)	o-o-o	o-v-o

CCC-CVC	risingCVC v falling	<i>mnʕ</i> (prohibit/forbid)	<i>muʕ</i> (swim)	s-s-o	s-v-o
CCC-CVC	risingCVC v falling	<i>rbħ</i> (win)	<i>rʕuħ</i> (go home)	s-o-o	s-v-o
CCC-CVC	risingCVC v falling	<i>nff</i> (scrape)	<i>nuf</i> (we are better)	s-o-s	s-v-s
CCC-CCC	rising v plateau	<i>zhd</i> (be strong)	<i>zld</i> (leather)	o-o-o	o-s-o
CCC-CCC	rising v plateau	<i>kff</i> (be faded)	<i>krf</i> (tie)	o-o-o	o-s-o
CCC-CCC	rising v plateau	<i>fth</i> (to operate)	<i>frʕh</i> (be happy)	o-o-o	o-s-o
CCC-CCC	rising v plateau	<i>zdm</i> (collect wood)	<i>zlm</i> (glance)	o-o-s	o-s-s
CCC-CCC	rising v plateau	<i>ħkm</i> (govern/judge)	<i>ħrʕm</i> (deprive)	o-o-s	o-s-s
CCC-CCC	falling v falling	<i>nsʕh</i> (advise)	<i>nʕħ</i> (pass a test)	s-o-o	s-o-o
CCC-CCC	falling v falling	<i>rdʕl</i> (borrow/lend)	<i>rgl</i> (lock)	s-o-s	s-o-s
CCC-CCC	falling v falling	<i>nkr</i> (to wake)	<i>ngr</i> (between)	s-o-s	s-o-s
CCC-CCC	falling v falling	<i>rdʕl</i> (borrow/lend)	<i>rʕhl</i> (leave the city)	s-o-s	s-o-s
CCC-CCC	falling v falling	<i>rgl</i> (lock)	<i>rʕhl</i> (leave the city)	s-o-s	s-o-s
CCC-CCC	plateau v plateau	<i>zʕbr</i> (prune)	<i>zgr</i> (cross)	o-o-s	o-o-s
CCC-CCC	plateau v plateau	<i>zbd</i> (to pull)	<i>zhd</i> (be strong)	o-o-o	o-o-o
CCC-CCC	plateau v plateau	<i>fkt</i> (give it)	<i>fst</i> (feed on)	o-o-o	o-o-o
CCC-CCC	plateau v plateau	<i>ʕbr</i> (disappear)	<i>ʕdr</i> (betray)	o-o-s	o-o-s
CCC-CCC	plateau v plateau	<i>bdr</i> (to mention)	<i>bsr</i> (spread)	o-o-s	o-o-s

Table X. Summary statistics for the *g*lmer on discrimination responses from Experiment 1.

<i>Output 1: MP Type Reference level = Non-MPs</i>	<i>Est</i>	<i>SE</i>	<i>z</i>	<i>p</i>
(Intercept)	2.09	0.45	4.61	<0.001
MP Type (CVC vs. CVC)	-1.67	0.55	-3.03	0.002
MP Type (CVC vs. CCC)	-0.96	0.47	-2.03	0.04
MP Type (CCC vs. CCC)	-1.10	0.46	-2.39	0.02
Style (Clear)	0.39	0.29	1.34	0.18
Lang (Tashlhiyt)	0.78	0.18	4.37	<0.001
MP Type (CVC vs. CVC)*Style (Clear)	0.96	0.34	2.79	0.005
MP Type (CVC vs. CCC)*Style (Clear)	0.13	0.30	0.42	0.67
MP Type (CCC vs. CCC)*Style (Clear)	0.38	0.30	1.28	0.20
Style (Clear) * Lang (Tashlhiyt)	0.59	0.21	2.75	.006
<i>Output 2: MP Type Reference level = CVC vs. CVC</i>				
(Intercept)	0.42	0.43	0.96	0.234
MP Type (CVC vs. CCC)	0.71	0.45	1.57	0.12

MP Type (CCC vs. CCC)	0.58	0.44	1.31	0.19
MP Type (Non-MPs)	1.67	0.55	3.03	0.002
Style (Clear)	1.35	0.24	5.72	<0.001
Lang (Tashlhiyt)	0.78	0.18	4.37	<0.001
MP Type (CVC vs. CCC)*Style (Clear)	-0.83	0.25	-3.31	<0.001
MP Type (CCC vs. CCC)*Style (Clear)	-0.57	0.24	-2.35	0.02
MP Type (Non-MPs)*Style (Clear)	-0.96	0.34	-2.79	0.01
Style (Clear) * Lang (Tashlhiyt)	0.59	0.21	2.75	0.01
<i>Output 3: MP Type Reference level = CCC vs. CCC</i>				
(Intercept)	0.99	0.31	3.24	0.001
MP Type (CVC vs. CVC)	-0.58	0.44	-1.31	0.19
MP Type (CVC vs. CCC)	0.14	0.33	0.41	0.68
MP Type (Non-MPs)	1.10	0.46	2.39	0.02
Style (Clear)	0.78	0.17	4.67	<0.001
Lang (Tashlhiyt)	0.78	0.18	4.37	<0.001
MP Type (CVC vs. CVC)*Style (Clear)	0.57	0.24	2.35	0.02
MP Type (CVC vs. CCC)*Style (Clear)	-0.25	0.19	-1.36	0.17
MP Type (Non-MPs)*Style (Clear)	-0.38	0.30	-1.28	0.20
Style (Clear) * Lang (Tashlhiyt)	0.59	0.21	2.75	0.01
<i>Num. observations = 5,328, Num. subjects = 72, Num. trial types = 37, Num. orderings = 4</i>				
<i>Retained model syntax: MP Type * Style + Style * Language Background (1+ Style Listener) + (1 Trial Type)+(1 Order)</i>				

Table X. Summary statistics for the post-hoc *glmer* on discrimination responses for trials containing CVC vs. CCC minimal pair comparisons from Experiment 1.

	<i>Est</i>	<i>SE</i>	<i>z</i>	<i>p</i>
(Intercept)	0.70	0.40	1.75	0.08

MP Type (CVC vs. CCC-Plateau)	0.84	0.47	1.78	0.08
MP Type (CVC vs. CCC-Falling)	0.11	0.53	0.21	0.84
Style (Clear)	0.52	0.22	2.30	0.02
Lang (Tashlhiyt)	0.87	0.22	4.05	<0.001
MP Type (CVC vs. CCC-Plateau) * Style (Clear)	-0.01	0.31	-0.05	0.96
MP Type (CVC vs. CCC-Falling) * Style (Clear)	0.84	0.35	2.38	0.02
<i>Num. observations = 1,728, Num. subjects = 72, Num. trial types = 12, Num. orderings = 4</i>				
<i>Retained model syntax: MP Type * Style + Language Background (1 + Style Listener) + (1 Trial Type)+(1 Order)</i>				

Table X. Summary statistics for the post-hoc *glmer* on discrimination responses for trials containing CCC vs. CCC minimal pair comparisons from Experiment 1.

	<i>Est</i>	<i>SE</i>	<i>z</i>	<i>p</i>
(Intercept)	1.82	0.50	3.66	<0.001
MP Type (CCC-Plateau vs. CCC-Plateau)	-0.94	0.62	-1.52	0.13
MP Type (CCC-Falling vs. CCC-Falling)	-1.36	0.62	-2.20	0.03
Style (Clear)	0.25	0.27	0.93	0.35
Lang (Tashlhiyt)	0.73	0.22	3.30	<0.011
MP Type (CCC-Plateau vs. CCC-Plateau) * Style (Clear)	0.84	0.33	2.55	0.01
MP Type (CCC-Falling vs. CCC-Falling) * Style (Clear)	0.51	0.31	1.61	0.11
Style (Clear) * Lang (Tashlhiyt)	0.59	0.28	2.11	0.03
<i>Num. observations = 2,160, Num. subjects = 72, Num. trial types = 15, Num. orderings = 4</i>				
<i>Retained model syntax: MP Type * Style + Style * Language Background (1 + Style Listener) + (1 Trial Type)+(1 Order)</i>				

Table X. Summary statistics for the *lmer* on nonword acceptability ratings from Experiment 2.

	<i>Est</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
(Intercept)	-0.41	0.09	79.72	-4.78	<0.001
Sonority Value (centered)	0.07	0.05	63.45	1.29	0.20
Style (Clear)	0.26	0.07	67.56	3.72	<0.001
Lang (Tashlhiyt)	0.56	0.11	70.25	5.07	<0.001
Sonority * Style (Clear)	0.18	0.06	1319.91	2.95	0.003
Sonority * Lang (Tashlhiyt)	0.01	0.06	244.73	0.12	0.90

Style (Clear) * Lang (Tashlhiyt)	0.00	0.10	67.57	-0.05	0.96
Sonority * Style (Clear) * Lang (Tashlhiyt)	-0.39	0.09	1320.09	-4.53	<0.001
<i>Num. observations = 1,512, Num. subjects = 72, Num. items = 21</i>					
<i>Retained model syntax: Sonority Value * Style * Language Background (1+Style + Sonority Value Listener)</i>					

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.

Bates, D. M. (2015). lme4: mixed-effects modeling with R <https://cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf>.

Beddor, P. S. (2009). A Coarticulatory Path to Sound Change. *Language*, 85(4), 785-821.

Bell, A. (1970). Syllabic consonants. *Working papers on language universals* (Stanford University) 4. B1–B49.

Benkí, J. R. (2001). Place of articulation and first formant transition pattern both affect perception of voicing in English. *Journal of Phonetics*, 29(1), 1-22.

Bensoukas, K. (2001). Stem Forms in the Nontemplatic Morphology of Berber. Rabat: University Mohammed 5 Doctorat d'Etat Thesis.

Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology*, 26(1), 75-108.

Bessell, N. J. (1998). Local and non-local consonant–vowel interaction in Interior Salish. *Phonology*, 15(1), 1-40.

Best, C. T., Tyler, M., Bohn, O., & Munro, M. (2007). Nonnative and second-language speech perception. *Language experience in second language speech learning*, 13-34.

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.

Bohn, O. S. (2017). Cross-language and second language speech perception. *The handbook of psycholinguistics*, 213-239.

Bouferroum, O., & Boudraa, M. (2015). CV coarticulation, locus and locus equation perspective on the invariance issue involving Algerian Arabic consonants. *Journal of Phonetics*, 50, 120-135.

Boukous, A. (1987). Phonotactique et Domaines Prosodiques en Berbère. Doctorat d'Etat dissertation, Université Paris VIII- Vincennes, Saint-Denis.

- Boukous, A. (2009). *Phonologie de l'Amazighe*. Rabat: IRCAM publications.
- de Boysson-Bardies, B. (1999). *How language comes to children: from birth to two years* (M. DeBevoise, Trans.). London: Cambridge.
- Bradlow, A. R. (2002). Confluent talker-and listener-oriented forces in clear speech production. *Laboratory phonology, 7*, 241-273.
- Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America, 112*(1), 272-284.
- Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of memory and language, 89*, 68-86.
- Chen, X., Ridouane, R., & Hallé, P. (2022). Mandarin listeners' onset-cluster perception according to sonority profile. In *The 10th International Symposium on the Acquisition of Second Language Speech (New Sounds 2022)*.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston and Mary E. Beckman (eds.) *Papers in Laboratory Phonology I: Between the grammar and physics of speech*. Cambridge, CUP, 283–333.
- Cohn, M., Segedin, B. F., & Zellou, G. (2022). Acoustic-phonetic properties of Siri-and human-directed speech. *Journal of Phonetics, 90*, 101123.
- Coleman, J. (2001). The phonetics and phonology of Tashlhiyt Berber syllabic consonants. *Transactions of the philological Society, 99*(1), 29-64.
- Cutler, A., Butterfield, S., & Williams, J. N. (1987). The perceptual integrity of syllabic onsets. *Journal of Memory and Language, 26*(4), 406-418.
- Daland, R., Hayes, B., White, J., Garellek, M., Davis, A., & Norrmann, I. (2011). Explaining sonority projection effects. *Phonology, 28*(2), 197-234.
- Davidson, L. (2011). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 270.
- Davidson, L., & Roon, K. (2008). Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association, 38*(2), 137-165.
- Dell, F. and M. Elmedlaoui. (1985). Syllabic Consonants and Syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics, 7*, 105-130.
- Dell, F., & Elmedlaoui, M. (2012). *Syllables in Tashlhiyt Berber and in Moroccan Arabic* (Vol. 2). Springer Science & Business Media.

- Easterday, S. (2019). *Highly complex syllable structure: A typological and diachronic study*. Language Science Press.
- El Aissati, A., McQueen, J. M., & Cutler, A. (2012). Finding words in a language that allows words without vowels. *Cognition*, 124(1), 79-84.
- Embarki, M., Yeou, M., Guilleminot, C., & Al Maqtari, S. (2007). An acoustic study of coarticulation in Modern Standard Arabic and Dialectal Arabic: pharyngealized vs non-pharyngealized articulation. *ICPhS XVI*, 141-146.
- Foss, D. J., & Gernsbacher, M. A. (1983). Cracking the dual code: Toward a unitary model of phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, 22(6), 609-632.
- Fougeron, C., & Ridouane, R. (2008). On the phonetic implementation of syllabic consonants and vowel-less syllables in Tashlhiyt. *Estudios de Fonética Experimental*, 140-175.
- Fowler, C. A. (1984). Segmentation of coarticulated speech in perception. *Perception & Psychophysics*, 36(4), 359-368.
- Frisch, S. A., Large, N. R., Zawaydeh, B., & Pisoni, D. B. (2001). Emergent phonotactic generalizations in English and Arabic. *Typological studies in Language*, 45, 159-180.
- Gafos, A. I. (2002). A grammar of gestural coordination. *Natural language & linguistic theory*, 269-337.
- Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. *Phonology*, 23(3), 387-429.
- Harrington, J., Kleber, F., Reubold, U., Schiel, F., & Stevens, M. (2019). The phonetic basis of the origin and spread of sound change. In *The Routledge handbook of phonetics* (pp. 401-426). Routledge.
- Hermes, A., Mücke, D., & Auris, B. (2017). The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics*, 64, 127-144.
- Hermes, A., Ridouane, R., Mücke, D., & Grice, M. (2011). Gestural Coordination in Tashlhiyt Syllables. In *ICPhS* (pp. 859-862).
- Jebbour, A. (1996). Morphologie et Contraintes Prosodiques en Berbère (Tachelhit de Tiznit)- Analyse Linguistique et Traitement Automatique. Doctorat d'Etat thesis, Mohammed V University, Faculty of Letters and Human Sciences, Rabat.
- Jebbour, A. (1999). Syllable Weight and Syllable Nuclei in Tachelhit Berber of Tiznit. *Cahiers de Grammaire 24: Phonologie: théorie et variation*, 95-116.
- Kabak, B., & Maniwa, K. (2007). L2 perception of English fricatives in clear and conversational speech: The role of phonemic, phonetic, and acoustic factors. In *Proceedings of ICPhS* (Vol. 2007).
- Kang, K. H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, 124(6), 3909-3917.

- Kirby, J. P. (2014). Incipient tonogenesis in Phnom Penh Khmer: Acoustic and perceptual studies. *Journal of Phonetics*, 43, 69-85.
- Krause, J. C., & Braid, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, 115(1), 362-378.
- Lahrouchi, M. (2010). On the internal structure of Tashlhiyt Berber triconsonantal roots. *Linguistic Inquiry*, 41(2), 255-285.
- Lahrouchi, M. (2018). Syllable structure and vowel-zero alternations in Moroccan Arabic and Berber. In *The Routledge Handbook of African Linguistics*, Augustine Agwuele & Adams Bodomo (eds.), 168-180. Oxon, UK: Routledge.
- Lieberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied*, 68(8), 1.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling*, 403-439.
- van der Lugt, A. H. (2001). The use of sequential probabilities in the segmentation of speech. *Perception & Psychophysics*, 63(5), 811-823.
- Maddieson, I. (2013). Consonant Inventories In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology Available online at <http://wals.info>.
- Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, 14(3), 380-407.
- McLeod, S., van Doorn, J., & Reed, V. A. (2001). Normal Acquisition of Consonant Clusters. *American Journal of Speech-Language Pathology*, 10, 99-110.
- Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27(2), 338-352.
- Munson, B. (2001). Phonological Pattern Frequency and Speech Production in Adults and Children. *Journal of Speech, Language, and Hearing Research*, 44(1-4), 778-792.
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, 34(3), 191-243.
- Ohala, J. J. (1981). Speech timing as a tool in phonology. *Phonetica*, 38(1-3), 204-212.
- Ohala, J. J. (1993). Coarticulation and phonology. *Language and speech*, 36(2-3), 155-170.
- Ohala, J. & Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Language and its ecology: Essays in memory of Einar Haugen*, 100, 343.

- Parker, S. G. (2002). *Quantifying the sonority hierarchy*. University of Massachusetts Amherst.
- Phatak, S. A., & Allen, J. B. (2007). Consonant and vowel confusions in speech-weighted noise. *The Journal of the Acoustical Society of America*, 121(4), 2312-2326.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 28(1), 96-103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 29(4), 434-446.
- Pouplier, M., & Beňuš, Š. (2011). On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory phonology*, 2(2), 243-273.
- Ridouane, R. (2008). Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. *Phonology*, 25(2), 321-359.
- Ridouane, R. (2014). Tashlhiyt Berber. *Journal of the International Phonetic Association*, 44(2), 207-221.
- Ridouane, R., & Fougeron, C. (2011). Schwa elements in Tashlhiyt word-initial clusters. *Laboratory Phonology*, 2(2), 275-300.
- Roach, P., Sergeant, P., & Miller, D. (1992). Syllabic consonants at different speaking rates: A problem for automatic speech recognition. *Speech Communication*, 11(4-5), 475-479.
- Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, 19(4), 499-526.
- Seyfarth, S., Buz, E., & Jaeger, T. F. (2016). Dynamic hyperarticulation of coda voicing contrasts. *The Journal of the Acoustical Society of America*, 139(2), EL31-EL37.
- Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, 118(3), 1677-1688.
- Smiljanić, R., & Bradlow, A. R. (2011). Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness. *The Journal of the Acoustical Society of America*, 130(6), 4020-4031.
- Tabain, M., & Butcher, A. (1999). Stop consonants in Yanyuwa and Yindjibarndi: locus equation data. *Journal of Phonetics*, 27(4), 333-357.
- Tilsen, S. (2016). Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics*, 55, 53-77.
- Tilsen, S., Zec, D., Bjorndahl, C., Butler, B., L'Esperance, M. J., Fisher, A., Heimisdottir, L., Renwick, M. & Sanker, C. (2012). A cross-linguistic investigation of articulatory coordination in word-initial consonant clusters. *Cornell Working Papers in Phonetics and Phonology*, 2012, 51-81.

Tupper, P., Leung, K. W., Wang, Y., Jongman, A., & Sereno, J. A. (2021). The contrast between clear and plain speaking style for Mandarin tones. *The Journal of the Acoustical Society of America*, 150(6), 4464-4473.

Woods, K. J., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics*, 79, 2064-2072.

Wright, R. A. (1996). *Consonant clusters and cue preservation in Tsou*. University of California, Los Angeles.

Wright, R. (2004). A review of perceptual cues and cue robustness. *Phonetically based phonology*, 34, 57.

Zec, D. (1995). Sonority constraints on syllable structure. *Phonology*, 12(1), 85-129.

Zellou, G., Lahrouchi, M., & Bensoukas, K. (2022). Clear speech in Tashlhiyt Berber: The perception of typologically uncommon word-initial contrasts by native and naive listeners. *The Journal of the Acoustical Society of America*, 152(6), 3429-3443.