Open Library of Humanities

The perception of vowelless words in Tashlhiyt

Georgia Zellou, Linguistics, University of California, Davis, CA USA, gzellou@ucdavis.edu

Mohamed Lahrouchi, CNRS & Université Paris 8, Paris, France, mlahrouchi@univ-paris8.fr

Karim Bensoukas, Mohammed V University in Rabat – FLHS, Rabat, Morocco, k.bensoukas@gmail.com

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 Tashlhiyt and non-native (English-speaking) listeners completed a paired discrimination task where the middle segment of the different-pair was either a vowel (e.g., *fan* vs. *fin*), consonant (e.g., *ʁbr* vs. *ʁdr*), or vowelless vs. voweled contrast (e.g., *tlf* vs. *tuf*). Experiment 2 was a word-likeness ratings task of tri-segmental nonwords constructed to vary in the sonority of the middle segment. We find that vowelless words containing different types of sonority profiles were generally highly discriminable by both native and non-native 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. At the same time, word-likeness ratings of nonwords indicated that listeners relied on their native-language experience to judge the wellformedness of new words: Tashlhiyt listeners were most likely to accept obstruent-centered vowelless words; meanwhile, English listeners' preferences increased with higher sonority values of the word center. Across both experiments, speech style variation provided further evidence as to how the phonetic implementation of vowelless words makes them perceptually stable. 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.

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 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 forms can shed light on some of the fundamental questions about how sound patterns emerge and evolve in linguistic systems. Tashlhiyt is also 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 (Fougeron & Ridouane 2008; Ridouane 2008; Ridouane & Fougeron 2011; Slovak in Pouplier & Beňuš 2011). However, 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). An examination of 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 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 and non-native (English-speaking) Tashlhiyt 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, tˤ, d, dˤ, n, r, rˤ, s, sˤ, z, zˤ, l, lˤ, ʃ, ʃˤ, ʒ, ʒˤ, j, k, kʷ, g, gʷ, w, q, qʷ, χ, χʷ, ʁ, ʁʷ, 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, such as word-initial sequences containing plateauing *kti*[1] or falling *rku* sonority profiles (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' (Ridouane 2008; Dell & Elmedlouai 2012; 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). However, 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 1978), 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; Bensoukas 2001; Ridouane 2008; Dell & Elmedlaoui 2012; 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 the 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 Lahrouchi's (2010) database were vowelless words containing an initial obstruent-sonorant sequence (rising sonority, e.g., *krz* 'plow'), 17% contained an initial cluster of obstruents (plateauing sonority; e.g., *gzm* 'cut'), 15% contained an initial sonorant-obstruent sequence (falling sonority; e.g., *rgl* 'knock'). 33% contained vowels (e.g. *knu* 'lean') and just 2% contained plateauing sonority with initial sonorants (SSO or SSS). Thus, within tri-segmental vowelless words in Tashlhiyt, rising (OSC; S = sonorant, O = obstruent, C = any consonant),[2] falling (SOC), and plateauing (OOC) sonority profiles are all frequent forms. (Note that the morphology of the language is sensitive to the sonority structure of consonant-only verbs. A case in point is imperfective internal vs. external gemination, as in the monosyllabic

---

[1] All Tashlhiyt forms provided are phonemic, unless otherwise noted in brackets.

[2] We use this simplified classification of consonant sonority to be consistent with Lahrouchi (2010), though we acknowledge that Dell and Elmedlaoui suggest using a more fine-grained sonority scale.

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 & Cooper-Leavitt 2019) 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 (Hermes et al. 2011; Ridouane et al. 2014; Hermes et al. 2017). Tashlhiyt listeners have been shown to treat consonant sequences as 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 (Hall 2006; Kirby 2014). 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 (Hall 2006; Davidson & Roon 2008). The presence of these variable vocoid elements in consonant sequences in Tashlhiyt 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 (Ridouane 2008; Dell & Elmedlaoui 2012; 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 of 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 only 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). However, 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. In particular, 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 plateauing 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 plateauing (e.g., *kp*) onset clusters with and without vowel epenthesis. Discrimination performance was higher for both rising and falling sonority clusters, presumably because the acoustic difference between the consonants made 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 from speaker-to-listener and, thus, evolutionarily dispreferred.

On the other hand, as outlined above, prior work on vowelless words in Tashlhiyt has found 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 intervocalically, 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. Berent et al. (2009, inter alia) 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 > plateauing > 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 consider 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 & Braida 2004; Zellou & Scarborough 2019; Cohn et al. 2022). Also, the acoustic effects of clear speech have been shown to increase intelligibility for listeners (Picheny et al. 1985; Smiljanić & Bradlow 2011; Scarborough & Zellou 2013). While clear speech effects have been found across several languages (e.g., Smiljanić & Bradlow 2005 for Croatian; Kang & Guion 2008 for Korean; Tupper et al. 2021 for Mandarin), 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 with 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., Buz et al. 2016; Seyfarth 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 non-native 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. 2007; 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 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*, which suggests 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., *tsskʃftstt* '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 (41AX) 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 *ʁbr* vs. *ʁdr*). 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 word-likeness task requires listeners to compare the phonetic and phonological properties of stimuli to the characteristics of the words stored in their memory. It has been shown that nonwords with greater lexical support are rated as more word-like than nonwords with less (Munson 2001; Frisch et al. 2001). We predict that Tashlhiyt listeners will show greater perceptual sensitivity and higher word-likeness 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 > plateauing > 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 > plateauing). Moreover, differences across tasks can speak to low-level auditory processes (discrimination) vs. higher-level phonological mechanisms (ratings) in the perception of vowelless words.

Figure 1: Examples of Tashlhiyt words varying in their sonority profiles (V = vowel, G = glide, L = liquid, N = nasal, VDF = voiced fricative, VLF = voiceless fricative; VDS = voiced stop; VLS = voiceless stop; O = obstruent; S = sonorant; C = any consonant).

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.

Thus, the present study is a comprehensive investigation of the perception of words without vowels in Tashlhiyt. We examine lexical discrimination performance and nonword acceptability judgements of tri-segmental vowelless words containing different sonority profiles by native Tashlhiyt and English-speaking listeners. We also varied the speaking style of the words (clear vs. casual speech) to explore the effect of systematic phonetic variation on the perceptual patterns of Tashlhiyt words. We predict that the patterns across words, tasks, listeners, and speech style can inform theoretical understandings of the relationship between language-specific phonetic variation and phonotactic patterns, as well as how auditory processing mechanisms influence preferred word forms.

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, provided in Table A in the Appendix.

In addition to the items selected for the discrimination task, 14 picture-able nouns were selected (*ajjis* 'horse', *tafukt* 'sun', *afullus* 'rooster', *azˤalim* 'onion', *ajdi* 'dog', *izi* 'fly', *ajjur* 'moon', *aʁrˤum* 'bread', *anzˤar* 'rain', *idukan* 'slippers', *tafunast* 'cow', *ilm* 'skin', *tafruχt* '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 __ baḥra* '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 & Braida 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 ($t(227) = 0.3$, $p = 0.76$).

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

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

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. **Figure 2** illustrates examples of the variation in the realization of vocoid-like elements across vowelless words produced for the current study. Vowelless words can be produced with no intervening vocoid element, as in the production of the word *nʒħ* 'pass a test' provided in **Figure 2.A**. In **Figure 2.B**, the production of the word *ʒbd* 'pull' is realized with a short vowel-like element (transitional vocoids are conventionally transcribed as [@] by Tashlhiyt scholars (Dell & Elmedlaoui 2012: 16) that has weak amplitude and formant structure. The production of the word *zdm* 'collect wood' in **Figure 2.C** is a full vowel, with longer duration than the element in 2.B. and clear, high frequency formant structure and large amplitude.

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; this only occurred for one word: *bdr* 'mention'). 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, across speaking styles.[3] The percentages of those elements that are full schwas across conditions, are also provided, as well as the duration of the element when present.

We ran tests to examine whether the rates of either any vocoid or the full schwa differs across speaking styles. A logistic regression run on presence of vocoid (=1) or not (=0) revealed that the likelihood of a vowelless token being produced with a vocoid-like element is *not* different across clear and casual productions [*est.* $= 0.15$, $z = 0.76$, $p = 0.4$]. A separate logistic regression

---

[3] This analysis was only performed for the non-excluded stimulus items (see Table A in the Appendix).
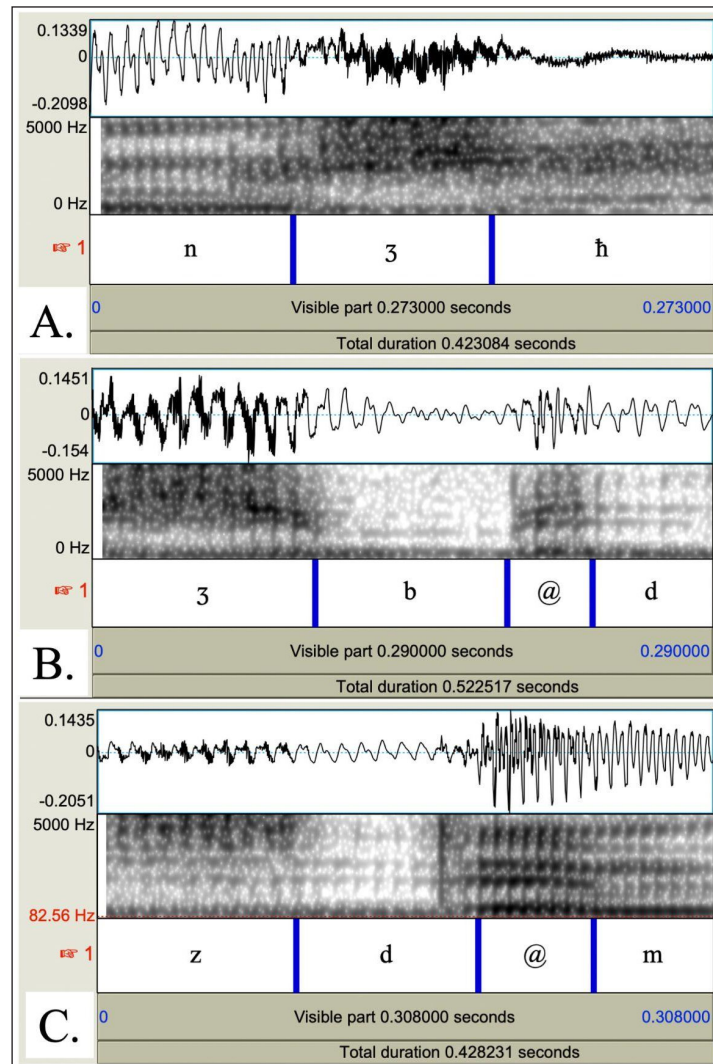
Figure 2: Waveform and spectrogram examples of vowelless words displaying no transitional element (*nʒħ* 'pass a test'; Panel A), a weak transitional vocoid element (*ʒbd* 'pull'; Panel B) and a transitional schwa vowel (*zdm* 'collect wood'; Panel C).

	% tokens with an inter-consonantal release or vocoid-like element	**% tokens with full vowel (subset of vocoid presence proportion)**	**Intrusive element duration (when present)**
Clear	60.7%	39.3%	.034 s (.05 s)
Casual	53.6%	26.8%	.019 s (.02 s)

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

run on presence of full schwa (=1) or not (=0) also reveals that the proportion of vowelless words with full vowels is *not* different across clear and casual productions [*est.* = 0.29, z = 1.4, p = 0.2]. However, a linear regression demonstrated that duration of the vocoid-like element, when present, is likely to be longer in clear speech productions [*est.* = 0.01, t = 2.1, p < 0.05]. Thus, vocoid-like elements, or full epenthetic schwas, are often produced in vowelless words (consistent with Ridouane & Fougeron 2011; Ridouane & Cooper-Leavitt 2019). 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)[4] 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 majority of Tashlhiyt participants were residing in Morocco at the time of the study (n = 24), while others were residing in France (n = 8), the US (n = 3), or UAE (n = 1). The English-speaking participants were recruited from the UC Davis Psychology 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

---

[4] One limitation of the language group comparison in the present study is that the native Tashlhiyt participants were older, on average, than the English-speaking participants. Future work making age-matched comparisons can explore whether age had an influence on the perception of vowelless words.

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 (Fowler 1984; Zellou 2017). 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 Table A in the Appendix. We noted after running the study that 4 of the critical trial pairs actually contain near-minimal pairs (marked with an asterisk in Table A). Those trials were excluded from the analysis reported in this paper. 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.

minimal pair type	Sonority Comparison (n of types)	Different Pair example	Same Pair example
non-MPs	non-minimal pairs (5)	*rˤuħ* vs. *sir*	*sir* vs. *sir*
CVC vs. CVC	risingCVC vs. risingCVC (5)	*fan* vs. *fin*	*fin* vs. *fin*
CVC vs. CCC	risingCVC vs. rising (4)	*tuf* vs. *tlf*	*tlf* vs. *tlf*
	risingCVC vs. plateau (5)	*fat* vs. *fkt*	*fkt* vs. *fkt*
	risingCVC vs. falling (3)	*nuf* vs. *nff*	*nff* vs. *nff*
CCC vs. CCC	rising vs. plateau (5)	*zlm* vs. *zdm*	*zdm* vs. *zdm*
	falling vs. falling (5)	*nsˤħ* vs. *nʒħ*	*nʒħ* vs. *nʒħ*
	plateau vs. plateau (5)	*ʁbr* vs. *ʁdr*	*ʁdr* vs. *ʁdr*

Table 3: Examples of the different paired discrimination trial types (full word list with glosses provided in Table A in the Appendix).

2.2. Results

2.2.1. Overall patterns

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. 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 = 3732.1, model ANOVA Chisq = 7.4, df = 1, $p < 0.01$) and was thus retained. Inclusion of the three-way interaction did not improve model fit. To explore significant interactions, *Tukey's HSD* pairwise comparisons were performed within the model, using the *emmeans()* function in the *emmeans* R package (Lenth et al. 2021).

Figure 3 shows mean discrimination performance for each Minimal Pair Type by Speaking Style, across listener groups. Even though there was not a 3-way interaction, **Figure 3** displays mean discrimination accuracy for each Minimal Pair Type across Speech Styles for each Listener Group so that the parallel patterns of performance for Tashlhiyt and English listeners can be
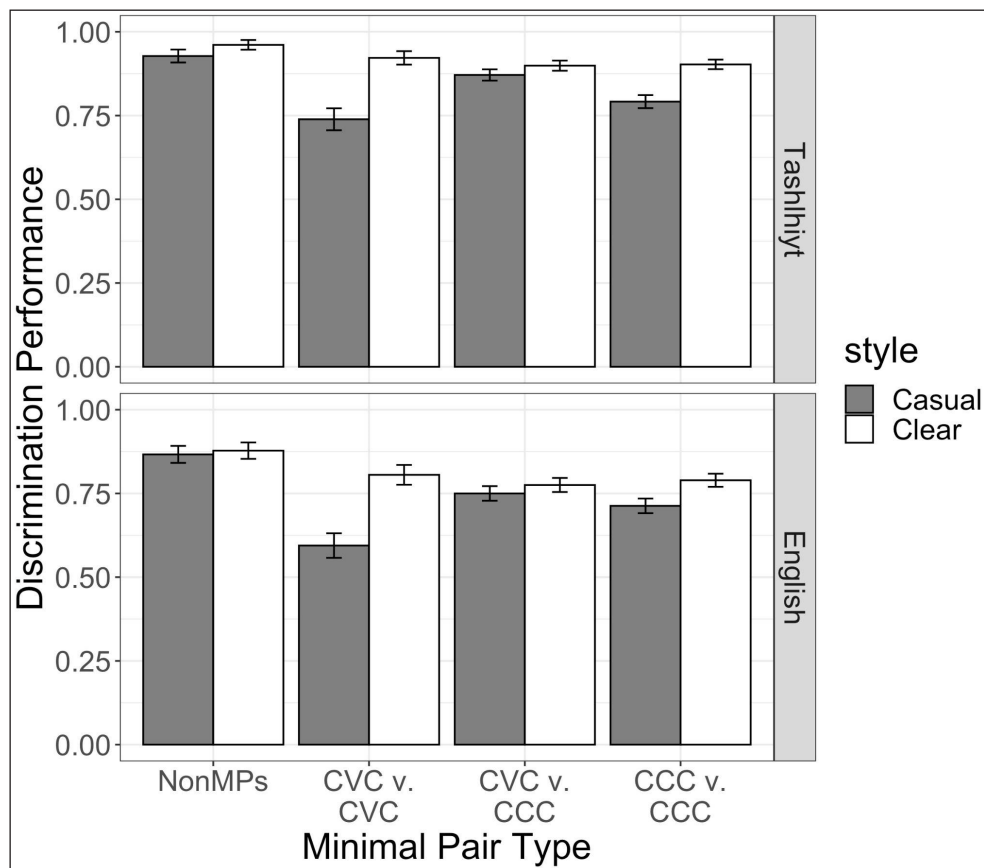

Figure 3: Experiment 1. Performance on paired discrimination trials across Minimal Pair and Speech Style conditions for native and non-native (English) Tashlhiyt listeners.

observed. The full model output is provided in Table B in the Appendix. The model revealed an effect of Minimal Pair Type. Discrimination performance was highest in trials where the different pair contained a non-minimal pair (91% accuracy) compared to trials containing voweled minimal pairs and vowelless minimal pairs (CCC vs. CCC: 80%; CVC vs. CVC: 77%). There was no difference between non-minimal pair trials and trials where the contrast was a voweled vs. vowelless minimal pair (CVC vs. CCC: 82%).

There was also an interaction between Minimal Pair Type and Speech Style. Post hoc *Tukey* tests were performed on the model with *emmeans* to unpack this interaction. First, the difference between discrimination performance for Casual and Clear speech styles was largest for the CVC vs. CVC comparison, and this difference was significant [*est.* = –1.6, z = –7.31, p < 0.001]. As seen in **Figure 3**, the lowest discrimination performance was for CVC vs. CVC trials in Casual speech. Post hoc tests also revealed that discrimination was worse in Casual speech for CCC vs. CCC trials [*est.* = –1.0, z = –6.4, p < 0.001]. However, there was no difference in discrimination across speech styles for non-minimal pairs [z = –2.3, p = 0.32] and for CVC vs. CCC trials [z = –2.8, p < 0.08]. Finally, a comparison between performance for non-minimal pair trials and CVC-CVC trials in Casual speech was significant, indicating that voweled trial discrimination was the lowest in reduced speech [*est.* = 1.7, z = 3.1, p < 0.05].

Language Background of the listener also predicted overall discrimination performance: Tashlhiyt listeners (87%) are better at discrimination of Tashlhiyt words than English-speaking listeners (77%). There was also an interaction between Language Background and Style: the positive coefficient value indicates that the difference between language backgrounds is greater in Clear speech than in Casual speech (indeed, the estimate for the post hoc comparison of non-native vs. native listeners for Clear speech [*est.* = –1.4, z = –4.6, p < 0.001] has a larger absolute value than that for the non-native vs. native Casual speech comparison [*est.* = –0.8, z = –3.9, p < 0.001]).

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.

2.2.2. Discrimination performance across sonority profiles

We also explore the extent to which sonority patterns *within* vowelless words lead to differences in perceptual sensitivity. To that end, two 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 the model with simple fixed effects, excluding all possible interactions, was computed as the best fit and thus retained (model ANOVA Chisq = 17.5, df = 1, $p < 0.001$). 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 Chisq = 4.8, df = 1, $p = .03$).

The summary statistics for both models are provided in Tables C and D the Appendix. The CVC vs. CCC model revealed effects of Speech Style (higher discrimination overall for Clear than Casual speech). There was also an effect of Minimal Pair Type. Performance was lowest for the CVC vs. CCC-Rising trials (75%), where sonority profile makes it harder to discriminate, compared to the CVC vs. CCC-Falling (88%) and CVC vs. CCC-Plateau trials (86%) which contain larger sonority differences. Speech style did not mediate these differences.

Within vowelless trials, performance was lowest when both pairs had Falling sonority (71%), compared to when both pairs had Plateauing sonority (78%); trials containing vowelless words with different sonority (Rising vs. Plateau) had the highest performance (87%). For the CCC vs. CCC model, there was a significant interaction between Minimal Pair Type and Speech Style, which can be seen in **Figure 4**. Post hoc *Tukey* tests revealed that while discrimination was lower in Casual than Clear speech for Plateau [*est.* = –1.5, z = –5.8, $p < 0.001$] and Falling [*est.* = –0.8, z = –3.1, $p < 0.05$] CCC minimal pairs, there was no casual speech reduction for Rising vs. Plateau vowelless words [$p = 0.33$]. Thus, vowelless pairs with matching sonority are harder to discriminate in reduced speech.

2.2.2.1. Post hoc investigating role of transitional vocoids

How might the realization of a transitional vocoid affect listeners' ability to discriminate vowelless words? We conducted a post hoc analysis to investigate this question. We tested whether the *duration* of the transitional vocoid present on the vowelless words in a different-pair affected listeners' discrimination performance for trials that contained the same sonority profile (plateau v. plateau and falling v. falling, since performance was lowest for them within CCC trials). To that end, we ran a mixed effects logistic regression model on the discrimination data for the plateauing and falling CCC vs. CCC trials. The model included fixed effects of longest Transitional Vocoid Duration (in seconds, continuous variable, centered and scaled; a value of 0 was coded for the few trials where there was no transitional vocoid on any of the words) and Minimal Pair
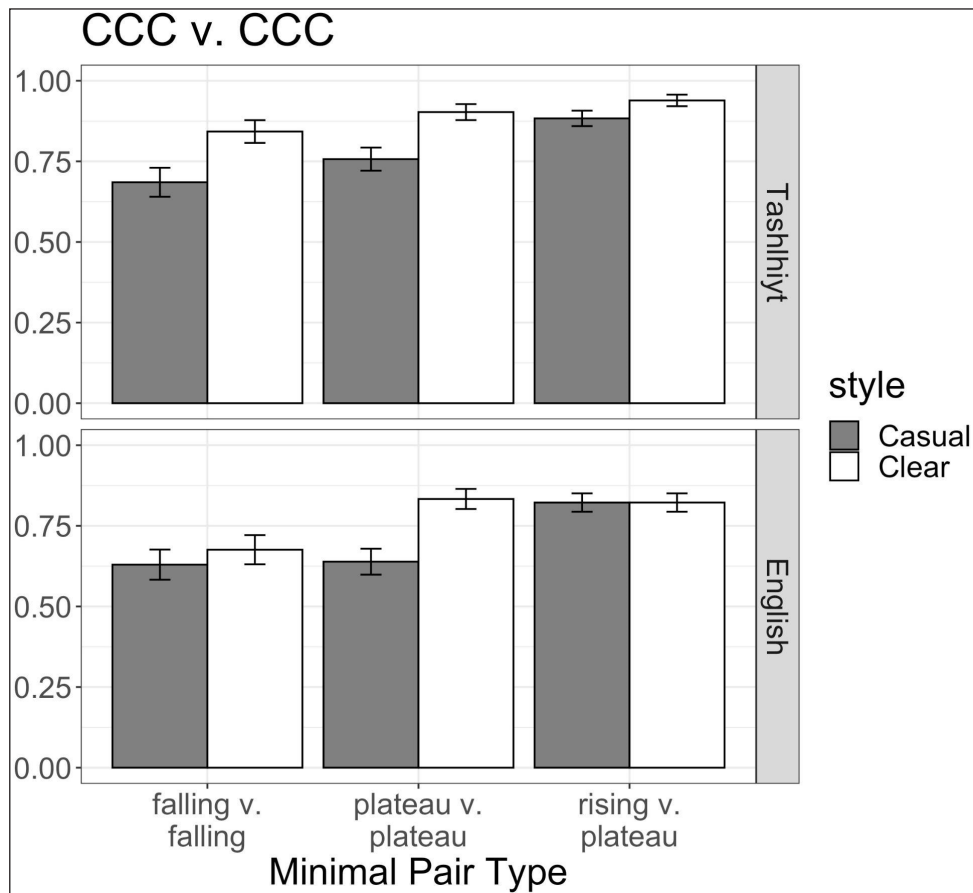
Figure 4: Experiment 1. Performance on paired discrimination trials with vowelless words only across Sonority Profile and Speech Style conditions for native Tashlhiyt and English listeners.

Type (Falling [reference level], Plateau), as well as their interaction. The model also included random intercepts for each participant, word pair type, and word pair ordering.

The model revealed a significant effect of Transitional Vocoid Duration [*est.* = 0.72, z = 2.5, $p < 0.05$] revealing that listeners perform better when the transitional element in a vowelless word is longer. There was also a significant interaction between Transitional Vocoid Duration and Minimal Pair Type indicating that longer vocoid duration provides an even larger boost for discrimination of plateauing sonority than for falling sonority vowelless words [*est.* = 1.04, z = 2.7, $p < 0.01$].

2.3 Interim discussion

In Experiment 1, native Tashlhiyt and non-native Tashlhiyt (English-speaking) listeners show variation in word 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 trials with words containing vowels. Thus, our results indicate that vowelless words are not perceptually harder than words with vowels.

We also found that vowelless word pairs containing falling sonority profiles (e.g., *nsˤħ* vs. *nʒħ)* are harder to discriminate than pairs containing plateaued sonority (e.g., *ʁbr* vs. *ʁdr*). 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).

Furthermore, a post hoc analysis revealed that listeners were better able to discriminate vowelless words when they contained longer transitional vocoids, and this was even stronger for plateauing sonority words. Thus, the phonetic realization of vowelless words in Tashlhiyt affects listeners' ability to perceive their internal structure.

Notably, while non-native listeners performed overall lower than native Tashlhiyt listeners, the perceptual discrimination patterns within each contrast were 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 as equally discriminable as other minimal pair trial types; in reduced speech, CVC minimal pairs are the hardest to discriminate). 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 (Smiljanić & Bradlow 2005; Kang & Guion 2008).

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, *rwl*). Thus, the non-words were constructed to be possible Tashlhiyt nonce words containing the most common C1s (*f, r*) and C3s (*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**. We note that an anonymous reviewer pointed out that the words containing two adjacent liquids (*flr, rln*) are phonotactically illicit in Tashlhiyt. Therefore, we expect nonwords *flr* and *rln* to receive very low ratings by the Tashlhiyt native listeners if word acceptability is guided by lexical statistics.

Nucleus	Nonword frame: /f_r/	Nonword frame: /r_n,m/
vowel	*fur*	*run*
liquid	*flr*	*rln*
nasal	*fnr*	*rmn*
voiced fricative	*fʐr*	*rʁm*
voiceless fricative	*fχr*	*rχn*
voiced stop	*fdr*	*rdn*
voiceless stop	*fqr*	*rtm*

Table 4: Nonwords analyzed for the Word Acceptability Ratings Task.

In addition to the 14 target nonwords analyzed in this study, 7 additional words were included in the experiment (*nul, nrl, nml, nʁl, nχl, ndl, nkl*). While these were originally intended to be nonwords, the use of the *n* as C1 is problematic as Tashlhiyt contains a productive first person plural prefix /n-/ as verbal inflection. Thus, some of our forms became real words (e.g., *nml* 'we show', *ndl* 'we cover', and *nkl* 'we spend the day') or could be interpreted by native listeners as an actual verb containing this prefix. Therefore, those forms were not included in the final analysis.

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 ƚéchelle de 0 à 100 dans quelle mesure il peut être un mot possible du tachelhit dans ƚavenir 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 of the 14 target nonword items 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.

The model output is provided in Table E in the Appendix. The model revealed an effect of Language Background wherein Tashlhiyt listeners provide higher acceptability ratings of the nonwords than English listeners (Tashlhiyt mean = 54.1, English mean = 36.1). There was also an effect of Speech Style: nonwords produced in Clear speech have higher ratings (Clear = 48.8, Casual = 41.4). 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 5**, 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
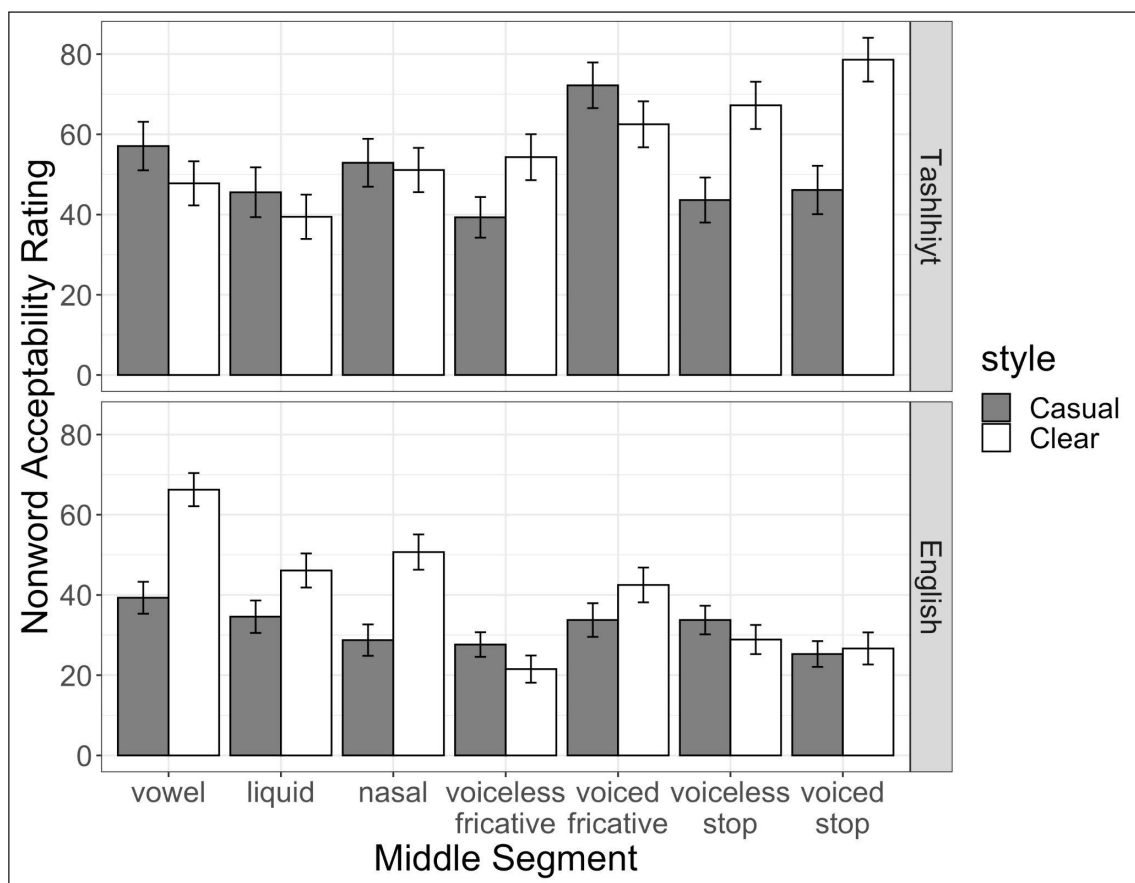

Figure 5: Nonword acceptability ratings by Middle Segment Type, by Speech Style for Tashlhiyt and English listeners.

across the board for the non-native Tashlhiyt 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 6** 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 6**: 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.
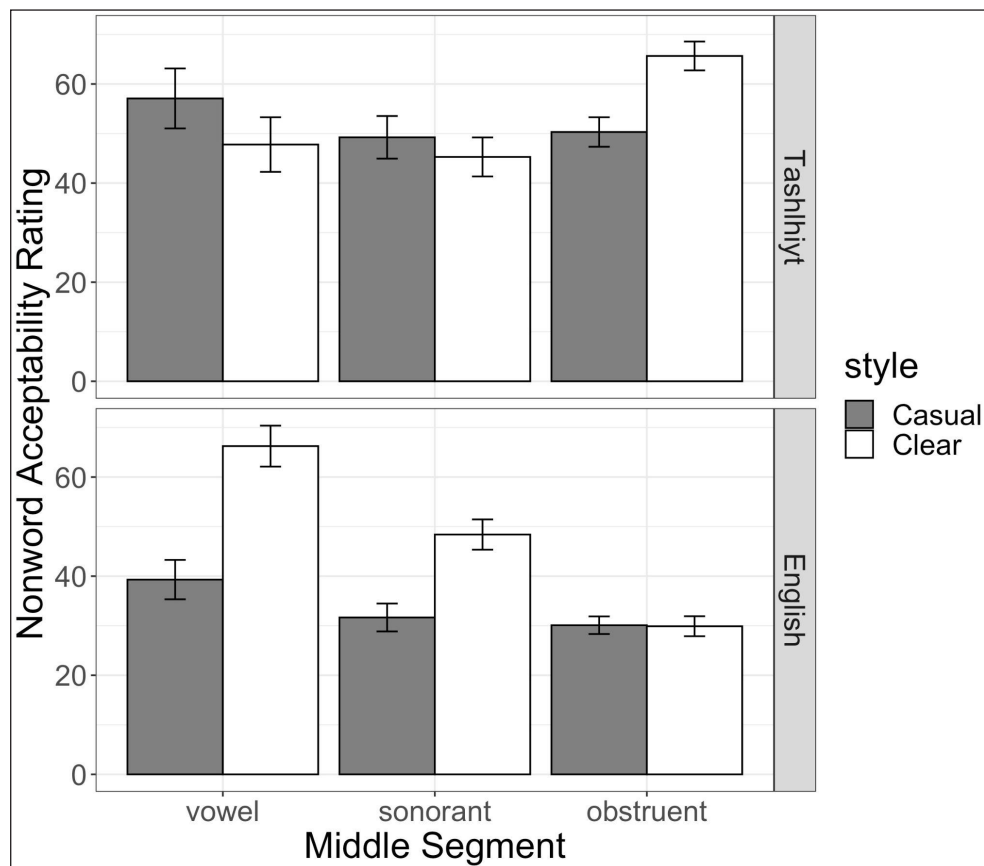
Figure 6: Nonword acceptability ratings where the middle segment is classified as a vowel, sonorant, or obstruent, by Speech Style for Tashlhiyt and English listeners.

4. General discussion

The current study investigated the perception of vowelless words in Tashlhiyt. Discrimination responses revealed that both native and non-native listeners are generally able to identify Tashlhiyt vowelless words contrasting in the middle consonant. However, Tashlhiyt- and English-speaking

listeners displayed distinct rating patterns of Tashlhiyt-like nonwords: native listeners were more likely to prefer obstruent-centered vowelless words, while non-native listeners preferred words with vowels. Across experiments, perceptual patterns were influenced by speaking style: in some cases, clear speech increased discriminability of and nonword acceptability ratings for vowelless words. In this discussion, we examine our results and consider their implications for phonological theory, phonotactic typology, as well as the relationship between the production and perception of speech.

4.1. Perception of vowelless words in Tashlhiyt

Vowelless words are rare in languages of the world. However, we find that they were not more difficult to discriminate in Tashlhiyt than words with vowels. This is surprising given that vowels provide robust transitional coarticulatory cues about the identity of adjacent sounds (Liberman 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 for speakers to enhance the acoustic cues to segment identity in vowelless words, even in casual speech where consonant cues are more likely to be obscured. Indeed, prior articulatory work has shown that consonant sequences are non-overlapping in Tashlhiyt and often produced with transitional vocoid elements that could provide acoustic cues to the segmental composition of vowelless words (e.g., Ridouane & Fougeron 2011). Indeed, the acoustic analysis of our stimuli revealed that vocoid-like elements, and even full schwas, are frequent in the vowelless items. Moreover, we found that they occur at similar rates across casual and clear speech, though they are longer in clear speech when they occur. And, we found that vowelless words containing longer transitional vocoid elements were easier to perceive. Thus, it appears that speakers do provide cues about the internal structure of vowelless words in Tashlhiyt. 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). In our study, these cues are present even when producing reduced speech, though they are enhanced in speech produced with the purpose of making words more intelligible. Our present findings are consistent with a *cue enhancement* view that vowelless words in Tashlhiyt are produced in such a way so as 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 observed that discrimination was harder when vowelless pairs contained matching word-onset sonority (here, falling vs. falling or plateau vs. plateau), than when sonority varied (rising vs. plateau). We interpret this as demonstrating that greater acoustic similarity across vowelless words made discrimination more difficult (Ohala & Kawasaki-Fukumori 1997). Moreover, discrimination performance for

sonority-matching pairs was even lower in casual speech. Thus, reduced speech made it even harder to perceive acoustically similar vowelless words. However, in clear speech, matching plateauing sonority items were not harder to discriminate than when vowelless pairs contained different sonority profiles. There was also a boost for the matching falling sonority items. Thus, clear speech did provide a boost for hard-to-discriminate vowelless contrasts.

Our acoustic analysis suggests that the *length* of the transitional elements is enhanced in clear speech. A post hoc analysis also established a relationship between the perceptual patterns and the realization of schwa across these word types: vowelless words are easier to discriminate when they contain transitional vocoids that are longer in duration. This provides a direct link between the phonetic properties of vowelless words to their perceptual stability. This also 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.

There are additional factors though, not considered in the present study, that might explain discrimination, or general perceptual, ability for vowelless words in Tashlhiyt. For instance, here, we focused only on the initial two segments of vowelless words. Yet, the phonological and phonetic properties of the word-final segment could impact phonetic and perceptual properties of vowelless words. Future work exploring the effect of the final consonant can be a fruitful direction. Moreover, the segmental composition of words has also been shown to predict vocoid realization. For instance, prior work has shown that schwa-like elements are common before final /r/ (Ridouane & Cooper-Leavitt 2019); many of our lexical items had a final /r/, therefore this could explain why some of our items with falling sonority were difficult to discriminate when they both contained a final /r/ (e.g., *nkr* vs. *ngr*). It could also explain why the rising vs. plateau comparisons were easier to discriminate when one pair contained an /r/ (e.g., *kʃʃ* vs. *krʃ*). Another consideration that we did not explore in the current study is that the laryngeal properties of adjacent consonants is another systematic factor conditioning the realization of vocoids within vowelless words. For instance, prior work has found that the occurrence of schwa-like elements relates to the voicing and place of articulation of adjacent consonants (e.g., Ridouane & Fougeron 2011). Also, pharyngealization of a segment within a vowelless word might influence its perceptibility (e.g., *ftħ* vs. *frˤħ*). Thus, there are indeed many additional factors that systematically predict the presence and duration of vocoid-like elements in vowelless words. These present ripe avenues for future work to explore their influence on the perception of these words in Tashlhiyt. Additionally, our classification of words in different sonority categories was based on a binary ranking of consonants (O vs. S). Work employing more fine-grained classifications of sonority could provide even further insights into how acoustic patterns of different segment types influence perceptual ability of Tashlhiyt vowelless words.

4.2. Perception of words with vowels in Tashlhiyt

We also observed that words containing vowels received the lowest discrimination performance in reduced speech. Why are voweled words in Tashlhiyt more difficult to perceive than vowelless words in this condition? One possibility is that strong consonant-to-vowel coarticulatory influences in Tashlhiyt leads 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 has been 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 that we observed in Tashlhiyt indicates greater consonant contrast preservation than vowel contrast 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 for lexical discrimination. Note, however, that the discrimination conditions containing voweled-vowelless minimal pairs (CVC vs. CCC) were equivalent in performance with the non-minimal pair trials, indicating the voweled and vowelless words were distinct enough from each other to be easy to distinguish perceptually. Within those trial types, words with similar acoustic profiles (CVC vs. CCC-Rising) were more difficult to discriminate in both reduced and clear speech, and those with the greatest acoustic difference in their sonority profile (e.g., CVC vs. CCC-Plateau) had higher performance across speech styles. Accordingly, Tashlhiyt vowels contrasting with a (obstruent) consonant were easier to discriminate, as opposed to vowels contrasting with sonorants and other vowels, which were 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 non-native listeners

We also compared native and non-native groups of listeners. Tashlhiyt listeners displayed overall higher discrimination performance than English listeners; however, the relative patterns across word types were 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 were robust differences in Tashlhiyt nonword acceptability ratings across listener groups. Tashlhiyt listeners were more accepting of vowelless nonwords with an obstruent center,

while English listeners preferred 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), so the obstruent-centered nonwords in Experiment 2 resemble many real words in Tashlhiyt. We also acknowledge that our liquid-centered nonwords actually violate Tashlhiyt phonotactics of having two adjacent liquids. Indeed, those forms received the lowest word-like ratings which is further support that listeners were recruiting their lexical knowledge (i.e., vowelless words with two adjacent liquids are not in their lexicon) when completing this task. Meanwhile, 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 was 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 were for the clear speech items. This is the first study, to our knowledge, comparing word-likeness ratings across items produced in different speaking styles. Presumably, the clear speech forms were easier to parse, phonologically, and thus provided more intelligible words that allowed 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 voweled nonwords as being more like a word of English.

More broadly, that we observed overall qualitative differences in the native and non-native 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 did not qualitatively change listeners' ability to discriminate when the word pairs were acoustically distinct enough. However, for the word-likeness task, performance recruits higher-level phonological processing mechanisms, and native-language-specific phonotactic patterns influenced 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 have 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 if the tongue configurations are incompatible (Ohala & Kawasaki-Fukumori 1997). Therefore, articulatory biases could explain why vowelless words are not more common cross-linguistically.

Moreover, as mentioned above, Tashlhiyt has a unique phonetic profile - 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 cross-linguistically 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 in languages where they are allowed. Language-specific phonetic implementation of sound sequences could explain the synchronic stability of complex phonotactic structures in Tashlhiyt. These speculations open lines of inquiry that can be explored in future work. For instance, future work directly comparing the perception of consonant clusters across languages that are more coarticulating (e.g., English) vs. less coarticulating (like Tashlhiyt) could 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, native and non-native listeners can discriminate between vowelless words. Yet, our results from a word-likeness task revealed 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.

Appendix

Word Type	Sonority Comparison	Word 1	Word 2	Word 1	Word 2
non-mps	non-mps	*rˤuħ* (go home)	*sir* (go!)	c-v-c	c-v-c
non-mps	non-mps	*luħ* (throw)	*ran* (they want)	c-v-c	c-v-c
non-mps	non-mps	*zˤurˤ* (visit)	*sak* (pass through)	c-v-c	c-v-c
non-mps	non-mps	*sul* (stay alive)	*mit* (what)	c-v-c	c-v-c
non-mps	non-mps	*zud* (like, as)	*liʁ* (I married)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	*fan* (they gave)	*fin* (they suppurated)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	*man* (which)	*mun* (accompany someone)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	*ʁar* (only)	*ʁir* (only)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	*lʒdid* (new)	*lʒdud* (ancestors)	c-v-c	c-v-c
CVC-CVC	risingCVC v risingCVC	*lfal* (omen)	*lfil* (elephant)	c-v-c	c-v-c
CCC-CVC	risingCVC v rising	*frˤħ* (be happy)	*fuħ* (revel in)	o-s-o	o-v-o
CCC-CVC	risingCVC v rising	*ʕlf* (feed)	*ʕif* (get tired of)	o-s-o	o-v-o
CCC-CVC	risingCVC v rising	*tlf* (get mixed up)	*tuf* (she's better)	o-s-o	o-v-o
CCC-CVC	risingCVC v rising	*slt* (leave on the sly)	*sut* (drink it)	o-s-o	o-v-o
CCC-CVC	risingCVC v plateau	*ʁdr* (betray)	*ʁar* (only)	o-o-s	o-v-s
CCC-CVC	risingCVC v plateau	*ʁbr* (disappear)	*ʁar* (only)	o-o-s	o-v-s
CCC-CVC	risingCVC v plateau	*fkt* (give it)	*fat* (give-2MS.PL)	o-o-o	o-v-o
CCC-CVC	risingCVC v plateau	*ftħ* (to operate)	*fuħ* (brag/boast)	o-o-o	o-v-o
CCC-CVC	risingCVC v plateau	*dbʁ* (tan)	*daʁ* (again)	o-o-o	o-v-o
CCC-CVC	risingCVC v falling	*mnʕ* (prohibit/forbid)	*muʕ* (swim)	s-s-o	s-v-o
CCC-CVC	risingCVC v falling*	*rbħ* (win)	*rˤuħ* (go home)	s-o-o	s-v-o
CCC-CVC	risingCVC v falling	*nʃf* (scrape)	*nuf* (we are better)	s-o-s	s-v-s
CCC-CCC	rising v plateau	*ʒhd* (be strong)	*ʒld* (leather)	o-o-o	o-s-o
CCC-CCC	rising v plateau	*kʃf* (be faded)	*krf* (tie)	o-o-o	o-s-o
CCC-CCC	rising v plateau	*ftħ* (operate)	*frˤħ* (be happy)	o-o-o	o-s-o
CCC-CCC	rising v plateau	*zdm* (collect wood)	*zlm* (glance)	o-o-s	o-s-s

Word Type	Sonority Comparison	Word 1	Word 2	Word 1	Word 2
CCC-CCC	rising v plateau	*ħkm* (govern/judge)	*ħrˤm* (deprive)	o-o-s	o-s-s
CCC-CCC	falling v falling	*nsˤħ* (advise)	*nʒħ* (pass a test)	s-o-o	s-o-o
CCC-CCC	falling v falling	*rdˤl* (borrow/lend)	*rgl* (lock)	s-o-s	s-o-s
CCC-CCC	falling v falling	*nkr* (wake)	*ngr* (between)	s-o-s	s-o-s
CCC-CCC	falling v falling*	*rdˤl* (borrow/lend)	*rˤħl* (leave the city)	s-o-s	s-o-s
CCC-CCC	falling v falling*	*rgl* (lock)	*rˤħl* (leave the city)	s-o-s	s-o-s
CCC-CCC	plateau v plateau*	*zˤbr* (prune)	*zgr* (cross)	o-o-s	o-o-s
CCC-CCC	plateau v plateau	*ʒbd* (pull)	*ʒħd* (be strong)	o-o-o	o-o-o
CCC-CCC	plateau v plateau	*fkt* (give it)	*fst* (feed on)	o-o-o	o-o-o
CCC-CCC	plateau v plateau	*ʁbr* (disappear)	*ʁdr* (betray)	o-o-s	o-o-s
CCC-CCC	plateau v plateau	*bdr* (mention)	*bsr* (spread)	o-o-s	o-o-s

Table A. Discrimination trial pairings. Trials marked with an asterisk contain near-minimal pairs and were excluded from the analysis.

	Est	SE	z	p
(Intercept)	2.11	0.44	4.77	<0.001
MP Type (CVC vs. CVC)	**–1.68**	**0.54**	**–3.12**	**0.001**
MP Type (CVC vs. CCC)	–0.72	0.47	–1.54	0.12
MP Type (CCC vs. CCC)	**–1.04**	**0.46**	**–2.26**	**0.02**
Speech Style (Clear)	0.35	0.29	1.20	0.23
Listener Language Background (Tashlhiyt)	**0.77**	**0.19**	**3.99**	**<0.001**
MP Type (CVC vs. CVC)*Style (Clear)	**0.97**	**0.34**	**2.85**	**0.004**
MP Type (CVC vs. CCC)*Style (Clear)	–0.18	0.31	–0.59	0.56
MP Type (CCC vs. CCC)*Style (Clear)	0.36	0.31	1.18	0.24
Style (Clear) * Lang (Tashlhiyt)	**0.59**	**0.21**	**2.79**	**.005**

Table B: Summary statistics for the *glmer* on discrimination responses from Experiment 1. *Num. observations = 4,752, Num. subjects = 72, Num. trial types = 33, Num. orderings = 4 Retained model syntax:* MP Type * Style + Style * Language Background (1+ Style | Listener) + (1|Trial Type)+ (1|Order).

	Est	SE	z	p
(Intercept)	0.70	0.34	2.05	0.04
MP Type (CVC vs. CCC-Plateau)	**0.83**	**0.38**	**2.20**	**0.02**
MP Type (CVC vs. CCC-Falling)	**1.03**	**0.50**	**2.07**	**0.03**
Speech Style (Clear)	**0.32**	**0.16**	**1.99**	**0.04**
Listener Language Background (Tashlhiyt)	**1.05**	**0.24**	**4.32**	**<0.001**

Table C: Summary statistics for the post-hoc *glmer* on discrimination responses for trials containing CVC vs. CCC minimal pair comparisons from Experiment 1.
Num. observations = 1,584, Num. subjects = 72, Num. trial types = 11, Num. orderings = 4
Retained model syntax: MP Type + Style + Language Background (1 + Style | Listener) + (1|Trial Type) + (1|Order).

	Est	SE	z	p
(Intercept)	1.90	0.50	3.77	< 0.001
MP Type (CCC-Plateau vs. CCC-Plateau)	–1.13	0.66	–1.71	0.09
MP Type (CCC-Falling vs. CCC-Falling)	–1.27	0.72	–1.77	0.08
Speech Style (Clear)	0.19	0.28	0.68	0.50
Listener Language Background (Tashlhiyt)	**0.57**	**0.25**	**2.33**	**0.02**
MP Type (CCC-Plateau vs. CCC-Plateau) * Style (Clear)	**0.99**	**0.34**	**2.89**	**< 0.01**
MP Type (CCC-Falling vs. CCC-Falling) * Style (Clear)	0.30	0.35	0.86	0.39
Style (Clear) * Lang (Tashlhiyt)	**0.67**	**0.31**	**2.20**	**0.03**

Table D: Summary statistics for the post-hoc *glmer* on discrimination responses for trials containing CCC vs. CCC minimal pair comparisons from Experiment 1.
Num. observations = 1,728, Num. subjects = 72, Num. trial types = 12, Num. orderings = 4
Retained model syntax: MP Type * Style + Style * Language Background (1 + Style | Listener) + (1|Trial Type) + (1|Order).

	Est	SE	df	t	p
(Intercept)	–0.41	0.09	79.72	–4.78	<0.001
Sonority Value (centered)	0.07	0.05	63.45	1.29	0.20
Speech Style (Clear)	**0.26**	**0.07**	**67.56**	**3.72**	**<0.001**
Listener Language Background (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**

Table E: Summary statistics for the *lmer* on nonword acceptability ratings from Experiment 2.
Num. observations = 1,008, Num. subjects = 72, Num. items = 14 Retained model syntax:
Sonority Value * Style * Language Background (1 + Style + Sonority Value | Listener).

Data availability statement

The data and R code for the project can be found at this link: https://osf.io/7bx4f/

Ethics and consent

Research was performed in accordance with the UC Davis Institutional Review Board (IRB number 1328085-2).

Acknowledgements

We thank Editor Björn Köhnlein and anonymous reviewers for their valuable feedback on this paper. Thanks also to those who participated in this study.

Competing interests

The authors have no competing interests to declare.

References

Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6), 716–723. DOI: https://doi.org/10.1109/TAC.1974.1100705

Bates, Douglas M. & Mächler, Martin & Bolker, Benjamin M. & Walker, Steven C. 2015. lme4: mixed-effects modeling with R https.cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf.

Beddor, Patrice S. 2009. A Coarticulatory Path to Sound Change. *Language* 85(4). 785–821. DOI: https://doi.org/10.1353/lan.0.0165

Bell, Alan. 1978. Syllabic consonants. In Joseph Greenberg (ed.), *Universals of human language, Vol. 2: Phonology*, 153–201. Stanford: Stanford University Press.

Benkí, José R. 2001. Place of articulation and first formant transition pattern both affect perception of voicing in English. *Journal of Phonetics* 29(1). 1–22. DOI: https://doi.org/10.1006/jpho.2000.0128

Bensoukas, Karim. 2001. *Stem Forms in the Nontemplatic Morphology of Berber*. Rabat, Morocco: University Mohammed V Doctoral Dissertation.

Berent, Iris & Lennertz, Tracy & Smolensky, Paul & Vaknin-Nusbaum, Vered. 2009. Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology* 26(1). 75–108. DOI: https://doi.org/10.1017/S0952675709001729

Bessell, Nicola J. 1998. Local and non-local consonant–vowel interaction in Interior Salish. *Phonology* 15(1), 1–40. DOI: https://doi.org/10.1017/S0952675798003510

Best, Catherine T. & Tyler, Michael & Bohn, Ocke-Schwen & Munro, Murray. 2007. Nonnative and second-language speech perception. *Language Experience in Second Language Speech Learning*. 13–34. DOI: https://doi.org/10.1075/lllt.17.07bes

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486357

Bohn, Ocke-Schwen. 2017. Cross-language and second language speech perception. In Traxler, Matthew & Gernsbacker, Morton Ann (eds.), *The Handbook of psycholinguistics*, 213–239.

Bouferroum, Othmane & Boudraa, Malika. 2015. CV coarticulation, locus and locus equation perspective on the invariance issue involving Algerian Arabic consonants. *Journal of Phonetics* 50. 120–135. DOI: https://doi.org/10.1016/j.wocn.2015.03.002

Boukous, Ahmed. 1987. *Phonotactique et Domaines Prosodiques en Berbère*. Paris, France: Université Paris VIII- Vincennes, Saint-Denis Doctoral Dissertation.

Boukous, Ahmed. 2009. *Phonologie de l'Amazighe*. Rabat: IRCAM publications.

Bradlow, Ann R. 2002. Confluent talker-and listener-oriented forces in clear speech production. *Laboratory Phonology* 7. 241–273. DOI: https://doi.org/10.1515/9783110197105.1.241

Bradlow, Ann R. & Bent, Tessa. 2002. The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America* 112(1). 272–284. DOI: https://doi.org/10.1121/1.1487837

Buz, Esteban & Tanenhaus, Michael K. & Jaeger, T. Florian. 2016. Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language* 89. 68–86. DOI: https://doi.org/10.1016/j.jml.2015.12.009

Chen, Xuejing & Ridouane, Rachid & Hallé, Pierre. 2022. Mandarin listeners' onset-cluster perception according to sonority profile. *The 10th International Symposium on the Acquisition of Second Language Speech (New Sounds 2022)*.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In Kingston, John & Beckman, Mary E. (eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech,* 283–333. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511627736.017

Cohn, Michelle & Ferenc Segedin, Bruno & Zellou, Georgia. 2022. Acoustic-phonetic properties of Siri- and human-directed speech. *Journal of Phonetics* 90. 101123. DOI: https://doi.org/10.1016/j.wocn.2021.101123

Coleman, John. 2001. The phonetics and phonology of Tashlhiyt Berber syllabic consonants. *Transactions of the Philological Society* 99(1). 29–64. DOI: https://doi.org/10.1111/1467-968X.00073

Cutler, Anne & Butterfield, Sally & Williams, John N. 1987. The perceptual integrity of syllabic onsets. *Journal of Memory and Language* 26(4). 406–418. DOI: https://doi.org/10.1016/0749-596X(87)90099-4

Daland, Robert & Hayes, Bruce & White, James & Garellek, Marc & Davis, Andrea & Norrmann, Ingrid. 2011. Explaining sonority projection effects. *Phonology* 28(2). 197–234. DOI: https://doi.org/10.1017/S0952675711000145

Davidson, Lisa. 2011. Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 37(1). 270–282. DOI: https://doi.org/10.1037/a0020988

Davidson, Lisa & Roon, Kevin. 2008. Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association* 38(2). 137–165. DOI: https://doi.org/10.1017/S0025100308003447

de Boysson-Bardies, Bénédicte. 1999. How language comes to children: from birth to two years (M. DeBevoise, Trans.). Cambridge, MA: MIT Press.

Dell, Francois & Elmedlaoui, Mohamed. 1985. Syllabic Consonants and Syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7. 105–130. DOI: https://doi.org/10.1515/jall.1985.7.2.105

Dell, Francois & Elmedlaoui, Mohamed. 2012. *Syllables in Tashlhiyt Berber and in Moroccan Arabic* (Vol. 2). Dordrecht: Springer Science & Business Media.

Easterday, Shelece. 2019. *Highly complex syllable structure: A typological and diachronic study*. Berlin: Language Science Press.

El Aissati, Abder & McQueen, James M. & Cutler, Anne. 2012. Finding words in a language that allows words without vowels. *Cognition* 124(1). 79–84. DOI: https://doi.org/10.1016/j.cognition.2012.03.006

Embarki, Mohamed & Yeou, Mohamed & Guilleminot, Christian & Al Maqtari, Sallal. 2007. An acoustic study of coarticulation in Modern Standard Arabic and Dialectal Arabic: pharyngealized vs non-pharyngealized articulation. *ICPhS XVI*. 141–146.

Foss, Donald J. & Gernsbacher, Morton A. 1983. Cracking the dual code: Toward a unitary model of phoneme identification. *Journal of Verbal Learning and Verbal Behavior* 22(6). 609–632. DOI: https://doi.org/10.1016/S0022-5371(83)90365-1

Fougeron, Cécile & Ridouane, Rachid. 2008. On the phonetic implementation of syllabic consonants and vowel-less syllables in Tashlhiyt. *Estudios de Fonética Experimental*, 140–175.

Fowler, Carol A. 1984. Segmentation of coarticulated speech in perception. *Perception & Psychophysics* 36(4). 359–368. DOI: https://doi.org/10.3758/BF03202790

Frisch, Stefan A. & Large, Nathan R. & Zawaydeh, Bushra & Pisoni, David B. 2001. Emergent phonotactic generalizations in English and Arabic. *Typological Studies in Language* 45. 159–180. DOI: https://doi.org/10.1075/tsl.45.09fri

Gafos, Adamantios I. 2002. A grammar of gestural coordination. *Natural Language & Linguistic Theory*, 269–337. DOI: https://doi.org/10.1023/A:1014942312445

Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23(3). 387–429. DOI: https://doi.org/10.1017/S0952675706000996

Harrington, Jonathan & Kleber, Felicitas & Reubold, Ulrich & Schiel, Florian & Stevens, Mary. 2019. The phonetic basis of the origin and spread of sound change. In Katz, William & Assmann, Peter (eds.), *The Routledge handbook of phonetics*, 401–426. New York, N.Y: Routledge. DOI: https://doi.org/10.4324/9780429056253-15

Hermes, Anne & Mücke, Doris & Auris, Bastian. 2017. The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics* 64. 127–144. DOI: https://doi.org/10.1016/j.wocn.2017.05.004

Hermes, Anne & Ridouane, Rachid & Mücke, Doris & Grice, Martine. 2011. Gestural Coordination in Tashlhiyt Syllables. *Proceedings of ICPhS* 2011. 859–862.

Jebbour, Abdelkrim. 1996. *Morphologie et Contraintes Prosodiques en Berbère (Tachelhit de Tiznit)-Analyse Linguistique et Traitement Automatique*. Rabat, Morocco: University Mohammed V Doctoral Dissertation.

Jebbour, Abdelkrim. 1999. Syllable Weight and Syllable Nuclei in Tachelhit Berber of Tiznit. *Cahiers de Grammaire* 24. 95–116.

Kabak, Bariş & Maniwa, Kazumi. 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, Kyoung-Ho & Guion, Susan 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, James P. 2014. Incipient tonogenesis in Phnom Penh Khmer: Acoustic and perceptual studies. *Journal of Phonetics* 43. 69–85. DOI: https://doi.org/10.1016/j.wocn.2014.02.001

Krause, Jean C. & Braida, Louis 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. DOI: https://doi.org/10.1121/1.1635842

Lahrouchi, Mohamed. 2010. On the internal structure of Tashlhiyt Berber triconsonantal roots. *Linguistic Inquiry* 41(2). 255–285. DOI: https://doi.org/10.1162/ling.2010.41.2.255

Lahrouchi, Mohamed. 2018. Syllable structure and vowel-zero alternations in Moroccan Arabic and Berber. In Agwuele, Augustine & Bodomo, Adams (eds.), *The Routledge Handbook of African Linguistics*, 168–180. New York, N.Y: Routledge. DOI: https://doi.org/10.4324/9781315392981-9

Lenth, Russell & Herve, Maxime & Love, Jonathon & Riebl, Hannes & Singman, Henrik. 2021. Package 'emmeans', [Software Package] (https://github.com/rvlenth/emmeans).

Liberman, Alvin M. & Delattre, Pierre C. & Cooper, Franklin S. & Gerstman, Louis J. 1954. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied* 68(8). 1–13. DOI: https://doi.org/10.1037/h0093673

Lindblom, Bjorn. 1990. Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling*, 403–439. DOI: https://doi.org/10.1007/978-94-009-2037-8_16

Maddieson, Ian. 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, Stefania & Pouplier, Marianne. 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. DOI: https://doi.org/10.1123/mcj.14.3.380

McLeod, Sharynne & van Doorn, Jan, & Reed, Vicki A. 2001. Normal Acquisition of Consonant Clusters. *American Journal of Speech-Language Pathology* 10. 99–110. DOI: https://doi.org/10.1044/1058-0360(2001/011)

Miller, George A. & Nicely, Patricia E. 1955. An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America* 27(2). 338–352. DOI: https://doi.org/10.1121/1.1907526

Munson, Benjamin. 2001. Phonological Pattern Frequency and Speech Production in Adults and Children. *Journal of Speech, Language, and Hearing Research* 44(1–4). 778–792. DOI: https://doi.org/10.1044/1092-4388(2001/061)

Norris, Dennis & McQueen, James M. & Cutler, Anne, & Butterfield, Sally. 1997. The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology* 34(3). 191–243. DOI: https://doi.org/10.1006/cogp.1997.0671

Ohala, John J. 1981. Speech timing as a tool in phonology. *Phonetica* 38(1–3). 204–212. DOI: https://doi.org/10.1159/000260024

Ohala, John J. 1993. Coarticulation and phonology. *Language and Speech* 36(2–3). 155–170. DOI: https://doi.org/10.1177/002383099303600303

Ohala, John & Kawasaki-Fukumori, Haruko. 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Language and its ecology: Essays in memory of Einar Haugen* 100. 343–365.

Parker, Stephen G. 2002. *Quantifying the sonority hierarchy*. Amherst, MA: University of Massachusetts Amherst Doctoral Dissertation.

Phatak, Sandeep A. & Allen, Jont B. 2007. Consonant and vowel confusions in speech-weighted noise. *The Journal of the Acoustical Society of America* 121(4). 2312–2326. DOI: https://doi.org/10.1121/1.2642397

Picheny, Michael A. & Durlach, Nathaniel I. & Braida, Louis 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. DOI: https://doi.org/10.1044/jshr.2801.96

Picheny, Michael A. & Durlach, Nathaniel I. & Braida, Louis 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. DOI: https://doi.org/10.1044/jshr.2904.434

Pouplier, Marianne & Beňuš, Štefan. 2011. On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory Phonology* 2(2). 243–273. DOI: https://doi.org/10.1515/labphon.2011.009

Ridouane, Rachid. 2008. Syllables without vowels: phonetic and phonological evidence from Tashlhiyt Berber. *Phonology* 25(2). 321–359. DOI: https://doi.org/10.1017/S0952675708001498

Ridouane, Rachid. 2014. Tashlhiyt Berber. *Journal of the International Phonetic Association* 44(2). 207–221. DOI: https://doi.org/10.1017/S0025100313000388

Ridouane, Rachid & Cooper-Leavitt, Jamison. 2019. A story of two schwas: a production study from Tashlhiyt. *Phonology* 36(3). 433–456. DOI: https://doi.org/10.1017/S0952675719000216

Ridouane, Rachid & Fougeron, Cécile. 2011. Schwa elements in Tashlhiyt word-initial clusters. *Laboratory Phonology* 2(2). 275–300. DOI: https://doi.org/10.1515/labphon.2011.010

Ridouane, Rachid & Hermes, Anne & Hallé, Pierre. 2014. Tashlhiyt's ban of complex syllable onsets: Phonetic and perceptual evidence. *STUF-Language Typology and Universals* 67(1). 7–20. DOI: https://doi.org/10.1515/stuf-2014-0002

Roach, Peter & Sergeant, Paul & Miller, Dave. 1992. Syllabic consonants at different speaking rates: A problem for automatic speech recognition. *Speech Communication* 11(4–5). 475–479. DOI: https://doi.org/10.1016/0167-6393(92)90054-B

Saltzman, Elliot & Byrd, Dani. 2000. Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science* 19(4). 499–526. DOI: https://doi.org/10.1016/S0167-9457(00)00030-0

Scarborough, Rebecca & Zellou, Georgia. 2013. Clarity in communication: "Clear" speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America* 134(5). 3793–3807. DOI: https://doi.org/10.1121/1.4824120

Seyfarth, Scott & Buz, Esteban & Jaeger, T. Florian. 2016. Dynamic hyperarticulation of coda voicing contrasts. *The Journal of the Acoustical Society of America* 139(2). EL31–EL37. DOI: https://doi.org/10.1121/1.4942544

Smiljanić, Rajka & Bradlow, Ann R. 2005. Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America* 118(3). 1677–1688. DOI: https://doi.org/10.1121/1.2000788

Smiljanić, Rajka & Bradlow, Ann 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. DOI: https://doi.org/10.1121/1.3652882

Tabain, Marija & Butcher, Andrew. 1999. Stop consonants in Yanyuwa and Yindjibarndi: locus equation data. *Journal of Phonetics* 27(4). 333–357. DOI: https://doi.org/10.1006/jpho.1999.0099

Tilsen, Sam. 2016. Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics* 55. 53–77. DOI: https://doi.org/10.1016/j.wocn.2015.11.005

Tilsen, Sam & Zec, Draga & Bjorndahl, Christina & Butler, Becky & L'Esperance, Marie-Josee & Fisher, Alison & Heimisdottir, Linda & Renwick, Margaret & Sanker, Chelsea. 2012. A cross-linguistic investigation of articulatory coordination in word-initial consonant clusters. *Cornell Working Papers in Phonetics and Phonology* 2012. 51–81.

Tupper, Paul & Leung, Keith W. & Wang, Yue & Jongman, Allard & Sereno, Joan 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. DOI: https://doi.org/10.1121/10.0009142

Van der Lugt, Arie H. 2001. The use of sequential probabilities in the segmentation of speech. *Perception & Psychophysics* 63(5). 811–823. DOI: https://doi.org/10.3758/BF03194440

Woods, Kevin J. & Siegel, Max H. & Traer, James & McDermott, Josh H. 2017. Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics* 79. 2064–2072. DOI: https://doi.org/10.3758/s13414-017-1361-2

Wright, Richard. 1996. *Consonant clusters and cue preservation in Tsou.* Los Angeles, CA: UCLA Doctoral Dissertation.

Wright, Richard. 2004. A review of perceptual cues and cue robustness. In Hayes, Bruce & Kirchner, Robert & Steriade, Donca (eds.), *Phonetically-based phonology,* 34–57. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486401.002

Zec, Draga. 1995. Sonority constraints on syllable structure. *Phonology* 12(1). 85–129. DOI: https://doi.org/10.1017/S0952675700002396

Zellou, Georgia. 2017. Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics* 61. 13–29. DOI: https://doi.org/10.1016/j.wocn.2016.12.002

Zellou, Georgia & Lahrouchi, Mohamed & Bensoukas, Karim. 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. DOI: https://doi.org/10.1121/10.0016579

Zellou, Georgia & Scarborough, Rebecca. 2019. Neighborhood-conditioned phonetic enhancement of an allophonic vowel split. *The Journal of the Acoustical Society of America* 145(6). 3675–3685. DOI: https://doi.org/10.1121/1.5113582