LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Segmentation of Highly Vocalic Speech Via Statistical Learning: Initial Results From Danish, Norwegian, and English

Fabio Trecca,[a] Stewart M. McCauley,[b] Sofie Riis Andersen,[a] Dorthe Bleses,[a] Hans Basbøll,[c] Anders Højen,[a] Thomas O. Madsen,[c] Ingeborg Sophie Bjønness Ribu,[d] and Morten H. Christiansen[a,e]

[a]Aarhus University, [b]University of Liverpool, [c]University of Southern Denmark, [d]University of Oslo, and [e]Cornell University

Research has shown that contoids (phonetically defined consonants) may provide more robust and reliable cues to syllable and word boundaries than vocoids (phonetically defined vowels). Recent studies of Danish, a language characterized by frequent long sequences of vocoids in speech, have suggested that the reduced occurrence of contoids may make speech in it intrinsically harder to segment than in closely related languages such as Norwegian. We addressed this hypothesis empirically in an artificial language learning experiment with native speakers of Danish, Norwegian, and English. We tested whether artificial speech consisting of concatenated contoid–vocoid syllables is easier to segment than speech consisting of vocoid–vocoid syllables where the first segment is a semivowel and the second a full vowel. Contrary to what was expected, we found

---

This research was partially supported by Danish Council for Independent Research (FKK) Grant DFF-7013-00074 awarded to Morten H. Christiansen. We are grateful for comments from Teresa Cadierno, Charlotte Gooskens, Padraic Monaghan, as well as from seven anonymous reviewers on an earlier version of this article.

This article has been awarded Open Materials and Open Data badges. All materials and data are publicly accessible via the Open Science Framework at https://osf.io/rmjtg. Learn more about the Open Practices badges from the Center for Open Science: https://osf.io/tvyxz/wiki.

Correspondence concerning this article should be addressed to Fabio Trecca, School of Communication and Culture, Aarhus University, Jens Chr. Schous Vej 2, DK-8000 Aarhus C, Denmark. E-mail: fabio@cc.au.dk

© 2018 Language Learning Research Club, University of Michigan
DOI: 10.1111/lang.12325

no effect of the phonetic makeup of the syllables on speech segmentability. Possible interpretations and implications of this result are discussed.

Keywords statistical learning; word segmentation; vowels; Danish; crosslinguistic; language distance

Introduction

Segmenting continuous speech into discrete meaningful units is a crucial step in early language acquisition for both first and second language learners (e.g., Carroll, 2004; Jusczyk, 1997). This task is not trivial because words without explicitly marked boundaries between them notoriously run together in speech. To succeed, children and adults make use of learning mechanisms that operate on a number of cues at the segmental, suprasegmental, and distributional level of the input (Curtin, Mintz, & Christiansen, 2005; Jusczyk, 1999; Thiessen & Saffran, 2003). For instance, both children and adults have been shown to be able to successfully segment fluent speech by tracking regularities such as transitional probabilities between syllables, a mechanism known as statistical learning (e.g., Erickson & Thiessen, 2015; Romberg & Saffran, 2010). At the same time, the availability and reliability of these segmentation cues can vary considerably across languages as a consequence of structural differences (Bates, Devescovi, & Wulfeck, 2001). As a result, languages may differ in their intrinsic degree of segmentability, with possibly crucial implications for how easily they are processed and acquired (cf. MacWhinney & Bates, 1989).

This issue has recently been raised by a number of studies of Danish, a North-Germanic language characterized by an unusually large number of vocoids compared to contoids in fluent speech (Basbøll, 2005). Vocoids are phonetically defined vowels: segments that have no constriction in the vocal tract, including full vowels, schwa vowels, semivowels/glides, and nonlateral approximants. By contrast, contoids are phonetically defined consonants: segments that result from the constriction of the airflow in the vocal tract, that is, obstruents, nasals, and lateral approximants. Pervasive diachronic and synchronic processes of lenition result in unusually long stretches of vocalic sounds that are realized as one single continuous vocoid with no clear amplitude or sound quality discontinuities, both within and across word boundaries. Examples of sentences that comprise only vocoids (e.g., *jeg er ude* [jɑ ɑ ˈuːð̞ə], "I'm out," or *er I ude?* [ˈæɐ̯ i ˈuːð̞ə], "are you out?") are common in Danish speech and can span as many as six syllables (Basbøll, 2012). Because of this unusually high rate of vocoids in speech, Danish has been suggested to be intrinsically harder to segment than other comparable languages (Bleses et al., 2008; Bleses & Basbøll, 2004; Bleses, Basbøll, Lum, & Vach, 2011). This suggestion has built on a number of previous studies (discussed below) suggesting that contoids, but not vocoids, provide the information over which segmentation is achieved (cf. Nespor, Peña, & Mehler, 2003).

In line with this hypothesis, there is both anecdotal and empirical evidence suggesting that Danish is particularly challenging to understand and learn for inexperienced listeners. For instance, a number of studies have found Danish to be hard to understand for speakers of Norwegian and Swedish, to both of which Danish is closely related genetically and typologically and both of which are reasonably intelligible for speakers of Danish (Gooskens, van Heuven, van Bezooijen, & Pacilly, 2010; Hilton, Schüppert, & Gooskens, 2011; Schüppert, Hilton, & Gooskens, 2016).

Other studies have suggested that Danish may also be a difficult language to acquire for children of Danish-speaking parents. Danish-learning children have been shown to lag behind in the early acquisition of receptive vocabulary between the ages of 8 and 15 months compared to children learning 12 other European and North American languages (Bleses et al., 2008) and in the acquisition of inflectional morphology up to 8 years of age compared to children from other Nordic countries (Bleses, Basbøll, & Vach, 2011; see also Bleses & Trecca, 2016; Kjærbæk, Christensen, & Basbøll, 2014). In a follow-up study to Bleses et al. (2008), crosslinguistic rates of vocabulary development were found to correlate negatively with the distribution of vocoids versus contoids in the respective languages, showing that the languages with a high ratio of vocoids to contoids in speech were associated with slower word learning rates, Danish having the highest ratio of vocoids and the slowest rate of lexical development of all the languages in the study (Bleses, Basbøll, Lum, et al., 2011). Last, recent experimental evidence from eye-tracking studies of Danish-learning children has also supported the idea of reduced segmentability in Danish speech (Trecca, Bleses, Højen, Madsen, & Christiansen, 2018; Trecca, Bleses, Madsen, & Christiansen, 2018). These studies have shown that long vocalic sequences that commonly occur in Danish child-directed speech (e.g., *her er . . .* [ˈheɐ̯ æɐ̯], “here is . . . ”) can impede children’s ability to identify a known object on screen when it is named as well as to learn novel nonsense words compared to when these familiar or unfamiliar words are presented in less heavily vocalic contexts.

Still, the mechanisms that make these sequences of vocoids in Danish speech putatively harder to process have not yet been investigated directly. The literature on the differential role of segments in speech has seemed to suggest that vocoids may inherently lack the information needed for statistical learning mechanisms to carry out the task of lexical identification (e.g., Bonatti, Peña, Nespor, & Mehler, 2005). This may be due to acoustic–phonetic properties of speech, making individual vocoids less distinctive, as well as to phonological–distributional aspects linked to the occurrence and arrangement of contoids versus vocoids in speech. These two aspects may be conflated, and we did not address the question of the origin of this putative bias in our study. Consequently, the more vocoids there are in speech, the less inherently segmentable the speech stream becomes because listeners would find it more difficult to derive useful information from vocoids than from contoids when tracking statistical regularities. In the case of Danish, this disadvantage may result in the observed processing and acquisition delay. We refer to this suggestion as the contoid-biased segmentation hypothesis.

However, a second possibility also seems plausible. By being particularly prone to coarticulation (van Ooijen, 1994) and to the assimilation/deletion of segments or even entire syllables (e.g., Wright, 2004), sequences of adjacent vocoids may make the speech signal particularly susceptible to degradation of the statistical (e.g., syllabic) structure of language over which statistical learning operates. That is, rather than attributing inherently lower segmentability to individual vocoids than to contoids (as in the contoid-biased segmentation hypothesis), we may instead posit that vocoids disrupt segmentation via statistical learning by making the statistical structure of language more diffuse and thus less informative. We refer to this second option as the syllable diffusion hypothesis. In our study, we approached this issue by explicitly testing the contoid-biased segmentation hypothesis in a crosslinguistic segmentation experiment.

Our categorization of segments as either contoids or vocoids—rather than as consonants and vowels—allows us to define segments based on their distinctive phonetic/articulatory features (i.e., whether each segment is produced, respectively, with or without any constriction of the airflow in the vocal tract) rather than on a phonological/functional distinction (i.e., based on whether segments form the sonority peak of a syllable or not; see Pike, 1943). Prototypically, vocoids coincide with syllable peaks and contoids with nonpeaks, although this is not always the case (such as the English nonsyllabic vocoids [j w] and the syllabic contoid [l̩]; see Basbøll, 2012); for additional information on the continuity between the articulatory and perceptual plan in the contoid–vocoid distinction, see Appendix S1 in the Supporting Information online.

Danish as a Prototypical Example of Highly Vocalic Speech Input

Among the European and North American languages, Danish has an unusually high ratio of vocoids to contoids (22 to 17, following Bleses, Basbøll, Lum, et al., 2011) both paradigmatically (as to its sound inventory) and syntagmatically (as to the sound structure of words and sentences in speech). This unusually high number of vocoids results from two factors. First, Danish has a large inventory of full vowels, counting 16 distinct vowel qualities in stressed syllables (13 of which have length contrasts) plus 49 possible diphthongs (Grønnum, 1998).[1] Second, most (earlier) obstruents occurring in the syllable-final unstressed position undergo pervasive lenition or weakening (Rischel, 1970/2009). Through lenition, these earlier obstruents in Danish lose their consonantal properties (e.g., closure) and become more sonorant, becoming in most cases vocoids (Bauer, 1983).

Together, these factors result in frequent long stretches of vocoids in speech, both within words—as in *(jeg) badede* [ˈb̥æːð̞əð̞ə], "(I) bathed," and across word boundaries, as in *(jeg) badede i åen* [ˈb̥æːð̞əð̞ə i ɔːˀn̩], "(I) bathed in the creek": The initial contoid [b̥] is followed by five vocoids in the former example, yielding a cv.vv.vv structure, where . marks a syllable boundary, and by as many as seven vocoids in the latter example, yielding a cv.vv.vv#v#vc structure, where # marks a word boundary (throughout the article, we use lowercase c and v to indicate contoids and vocoids, and uppercase C and V to indicate consonants and full vowels). This is in sharp contrast to closely related languages such as Norwegian, where the (interchangeable) cognate verbs *(jeg) badet* [ˈbɑːdət] or *(jeg) bada* [ˈbɑːdɑ], "(I) bathed," yield a cvc.vc or cvc.v structure, respectively (Ragnarsdóttir, Simonsen, & Plunkett, 1999). For more detailed accounts of the sound structure of Danish, we refer the reader to Basbøll (2005) and Grønnum (1998).

Are Vocoids Less Useful Than Contoids for Speech Segmentation?

As we stated earlier, a large body of literature has suggested that vocoids may be intrinsically less informative than contoids for segmentation. For instance, Mattys and Jusczyk (2001) found that English-learning children up to 16 months of age fail to segment vocoid-initial words in fluent speech while accurately segmenting contoid-initial words (e.g., they segment *dice* from both *roll#dice* and *cold#ice*). Analogous results were found by Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, and Jusczyk (2005), where English-learning children up to 13.5 months of age successfully segmented contoid-initial verbs from fluent speech but failed to segment vocoid-initial verbs. Elsewhere, children at a similar age were found to successfully segment vocoid-initial words, but only when these occurred in highly salient positions, for example, sentence initially (Seidl & Johnson, 2008) or when they were preceded by high-frequency function words such as *the* (Kim & Sundara, 2015). Moreover, there is considerable evidence that contoids may carry the most weight in lexical identification (Nespor et al., 2003). Studies on the differential role of consonants and vowels (here intended phonologically and functionally as segments that either do or do not constitute the peak of a syllable) in segmentation via statistical learning have shown that adults easily pick up on statistical regularities computed over tiers of Cs (e.g., *b_d_k_*), but not of Vs (e.g., *_a_e_o*), in strings of nonsense words consisting of CV syllables, for example, *badeko* (Bonatti et al., 2005; Mehler, Peña, Nespor, & Bonatti, 2006; but see Newport & Aslin, 2004). For this reason, phonological consonants have been often claimed to be crucial cues for segmentation, but phonological vowels have been suggested to be more useful for encoding syntactic, morphologic, and prosodic information (e.g., Gervain & Mehler, 2010).

In explaining this apparent primacy for contoids over vocoids, two explanations are possibly conflated. The first explanation is acoustic–phonetic and lies in the fact that vocoids create a virtually steady-state acoustic signal with no interruption because they are realized through the free flow of air in the vocal tract and therefore are highly sonorant (Stevens, 1998). Contoids, by contrast, add temporally salient perceptual discontinuities to the otherwise continuous speech signal (e.g., Lieberman, Harris, Hoffman, & Griffith, 1957). The alternation of contoids and vocoids in speech may therefore maximize the perceptual salience of syllables in the speech stream (e.g., Wright, 2004; see also Oller, 2000), which possibly also explains why cv syllables are predominant in canonical babbling and why children's early vocabulary mostly consists of words with cv, cvc, or cvcv structure (Stoel-Gammon, 1998, 2011; see also Gonzalez-Gomez, Hayashi, Tsuji, Mazuka, & Nazzi, 2014; MacNeilage & Davis, 2000).

At the same time, the bias in favor of contoids may stem from learned phonological regularities in the ambient language (e.g., Keidel, Jenison, Kluender, & Seidenberg, 2007). For example, contoids outnumber vocoids in most languages (e.g., Hochmann, Benavides-Varela, Nespor, & Mehler, 2011), and the majority of words in child-directed speech are contoid initial in most languages (e.g., in English: Swingley, 1999; in Danish: Bleses, Vach, Wehberg, Faber, & Madsen, 2007). This idea is supported by evidence that children learning, for example, French and Italian (e.g., Havy & Nazzi, 2009; Hochmann et al., 2011) show a bias for contoids when learning new words. This bias seems to emerge during the first year of life, superseding and eventually replacing entirely a bias for vocoids documented in very early infancy (e.g., Hochmann, Benavides-Varela, Fló, Nespor, & Mehler, 2017; Nishibayashi & Nazzi, 2016; see also Bouchon, Floccia, Fux, Adda-Decker, & Nazzi, 2015). Interestingly, a recent study by Højen and Nazzi (2016) found that Danish-learning children still show a bias in favor of vocoids in word learning at as late as 20 months of age, possibly because of the highly vocalic nature of their ambient language, raising the possibility that the primacy of contoids as segmentation cues may be partly language specific. Although our study did not address questions regarding the ontology of the putative segmentation bias for contoids, we acknowledge that a vocoid bias may better equip proficient speakers of Danish for segmenting highly vocalic speech compared to speakers of less vocoid-heavy languages. In our study, we controlled for these baseline differences by testing participants with different native languages.

It is important to note that many of the studies reviewed previously may pertain more to the distinction between consonants and vowels as phonological–functional elements rather than to the phonetic–perceptual distinction between contoids and vocoids. Despite the large overlap between the acoustic–phonetic categorization into contoids versus vocoids and the correlated phonological categorization into consonants versus vowels, our study only addressed the question pertaining to the former distinction directly because previous hypotheses about Danish have traditionally been concerned primarily with phonetic–perceptual aspects of speech. Possible issues with disentangling the phonetic versus phonological levels in our experimental design are discussed below.

The Present Study

On the basis of the literature that we reviewed above, we set out to test the contoid-biased segmentation hypothesis by using a crosslinguistic artificial language learning task with adult speakers of Danish, Norwegian, and English. We created two artificial speech streams (in the style of Saffran, Newport, & Aslin, 1996) with identical statistical structures but different phonetic makeup. The first speech stream consisted of repeated contoid–vocoid syllables (the cv language) whereas the second speech stream consisted of repeated vocoid–vocoid syllables (the vv language) in which the first vocoid was always realized as a semivowel derived from weakening the contoids from the cv language. We predicted that, if our hypothesis was correct, then our participants would be significantly better at learning words from the cv language than those from the vv language.

Part I: Preliminary Corpus Analyses

We ran some preliminary corpus analyses with the purpose of verifying the notion that different languages—in this case, Danish and English—differ with respect to the distribution of contoids and vocoids. To do this, we looked at the distribution of four diphone classes in Danish and English child-directed speech corpora: (a) contoid–contoid (cc), (b) vocoid–vocoid (vv), (c) contoid–vocoid (cv), and (d) vocoid–contoid (vc). The use of child-directed speech corpora was intended to mimic the language to which young as well as adult beginner learners of Danish might be exposed.

Method

We prepared four Danish corpora of child-directed speech: Anne (1;1–2;11) and Jens (1;0–3;1) from the Danish Plunkett corpus (Plunkett, 1985, 1986) in the Child Language Data Exchange System (CHILDES) database (MacWhinney, 2000) and Fam03 (0;9–2;6) and Fam05 (0;9–2;6) from the Odense Twin Corpus (Basbøll et al., 2002). Each Danish corpus was matched, according to size and vocabulary range, with an American English single-child corpus of child-directed speech from the CHILDES database. These included the Trevor (Demetras corpus; Demetras, 1989; 2;0–3;11), Brooklyn (Brent corpus; Brent & Cartwright, 1996; 0;8–1;2), Nathaniel (Snow corpus; MacWhinney & Snow, 1985; 2;5–3;9), and Joe (Soderstrom corpus; Soderstrom, Blossom, Foygel, & Morgan, 2008; 0;5–1;0) single-child corpora. We prepared each corpus using an automated procedure whereby markup and tags were removed except for word boundary markers. We then classified each phoneme type as either contoid or vocoid based on Basbøll (2005). Following corpus preparation, we tested the distribution of each classified phoneme type across the different corpora.

Results and Discussion

The key finding was that the distribution of diphone types looks markedly different for Danish versus English child-directed speech corpora. As can be seen in Figure 1, cc diphones occur at a much higher rate in English (38.9%) than in Danish (7.4%); vc diphones occur almost at the same rate in Danish (24.6%) and English (27.8%), and cv diphones occur at a higher rate in Danish (49.8%) than in English (27.3%). We were particularly interested in the rate of vv diphones (i.e., the type of diphone that would be hardest to segment according to our hypothesis). This was numerically higher in Danish (18.2%) than in English (6.1%), and 23% of all the Danish vv diphones versus 9% of the English ones contained a word boundary. Danish-learning children and adults

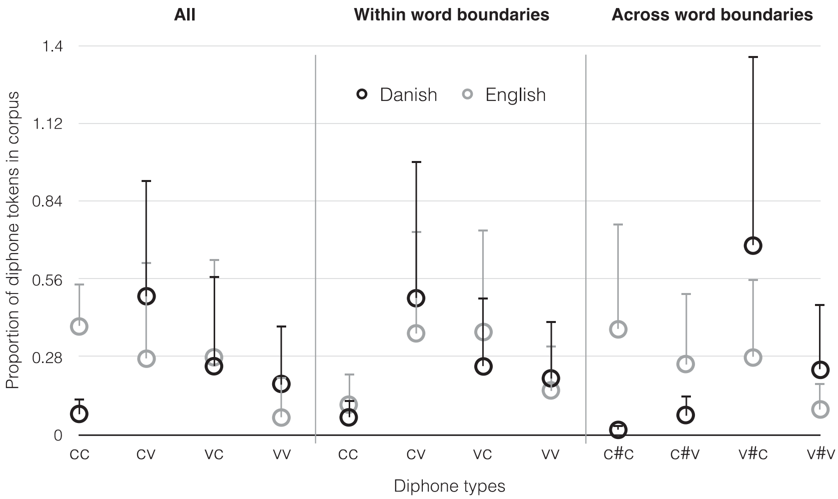

Figure 1 Distribution of diphone pairs by type (cc, cv, vc, and vv) in Danish and English child-directed speech corpora, with (v#v) or without (vv) a word boundary occurring within the pair, where c = contoid (i.e., obstruents, nasals, and lateral approximants) and v = vocoid (i.e., full vowels, schwa vowels, semivowels/glides, and nonlateral approximants).

are thus approximately three times more likely to encounter a word boundary between two vocoids in speech.

This result provided further motivation for our study by showing that vv diphones are indeed more frequent in Danish than in English, especially across word boundaries. This empirical finding lends some initial support to the notion that both first and second language learners of Danish may indeed face a more difficult task in segmenting the speech input relative to learners of other Germanic languages such as English and to the hypothesis that Danish-learning children's delayed acquisition may be partly rooted in structural properties of the language itself (e.g., Bleses, Basbøll, Lum, et al., 2011). Although the classification of segments into either contoids or vocoids provided us with information about the phonetic nature of Danish child-directed speech corpora, the corpora used in this study were in citation form and thus did not capture the pervasive segmental and syllabic reduction typical of spoken Danish. This possibly made our analyses more conservative by underestimating the number of word boundaries occurring within vv diphones in Danish.

Part II: Artificial Speech Segmentation Experiment

The purpose of our artificial speech segmentation experiment was to test whether speech consisting of alternating contoid–vocoid patterns, that is, the cv language, where the first segment was a consonant and the second a full vowel, is inherently easier to segment by way of statistical learning than speech consisting exclusively of vocoids, that is, the vv language, where the first segment is realized as a semivowel and the second segment is a full vowel. To do this, we used an artificial speech segmentation procedure with a sample of adult native speakers of Danish, Norwegian, and American English. Norwegian and English have different degrees of relatedness to Danish. Norwegian is especially close to Danish, both genetically and typologically, but Norwegian and English both differ from Danish in having a considerably less vocalic phonetic structure. The vocoid-to-contoid ratio in the three languages is 1.29 for Danish, 0.86 for Norwegian, and 0.67 for English (Bleses, Basbøll, Lum, et al., 2011).[2] Including participants with different native languages allowed us to control primarily for effects of different vocoid-to-contoid ratios in the ambient language and for those of familiarity with the phonetic inventory of the speech stimuli.

Method

Participants

We tested a total of 186 adult native speakers from three different native language groups: Danish ($n = 56$, $M_{\text{age}} = 25$ years, $SD = 6$, 48 females), Norwegian ($n = 56$, $M_{\text{age}} = 29$ years, $SD = 11$, 37 females), and American English ($n = 56$, $M_{\text{age}} = 20$ years, $SD = 1$, 42 females). Within each native language group, half of the participants ($n = 28$) listened to the cv language whereas the other half of the participants listened to the vv language. The participants were university students and staff from three universities: University of Southern Denmark (Odense, Denmark), University of Oslo (Oslo, Norway), and Cornell University (Ithaca, NY). All participants grew up in monolingual families. Danish and Norwegian participants did not receive compensation for their participation; American student participants received course credit.

Materials

Two distinct artificial languages, a cv language and a vv language, were created for use as training materials. Each language was generated by combining a set of 11 naturally spoken Danish phonemes into 15 syllables, then these were combined into the sets of three disyllabic and three trisyllabic nonsense words shown in Table 1.[3] These words were then joined together into two 11-minute-long streams with a speech rate of about 205 syllables per minute

Table 1 Words and nonword foils in the two artificial languages

	Words		Foils	
Structure	cv language	vv language	cv language	vv language
Disyllables	cv.cv	vv.vv	cv.cv	vv.vv
	/gʌdɔ/	/jʌðɔ/	/dəbu/	/ðewu/
	/dədʌ/	/ðeðʌ/	/bædɔ/	/wæðɔ/
	/dɛdɑ/	/ðɛðɑ/	/bɔdʌ/	/wɔðʌ/
Trisyllables	cv.cv.cv	vv.vv.vv	cv.cv.cv	vv.vv.vv
	/gubɔgɔ/	/juwɔjɔ/	/gʌdɑgə/	/jʌðɑjə/
	/bebugə/	/wewujə/	/bɑgædɛ/	/wɑjɛðɛ/
	/bægæbɑ/	/wæjɛwɑ/	/gɔgube/	/jɔjuwe/

Note. The phonetic structure of the words and foils is indicated by c = contoid (i.e., obstruents, nasals, and lateral approximants) and v = vocoid (i.e., full vowels, schwa vowels, semivowels/glides, and nonlateral approximants).

by repeating 150 tokens of each word in random order. Words never occurred twice in a row; however, six repetitions of each of the six word-final syllables in each language were added randomly to the chain as part of a cover task (see Procedure). The two artificial languages consisted of a series of equally stressed syllables, resulting in a staccato-sounding speech stream with no pauses or prosodic patterns. Figure 2 shows amplitude envelopes and spectrograms for excerpts of the two artificial languages. Although the staccato nature of the speech stream made it relatively undemanding for participants to discern each syllable boundary, the only information about word boundaries was provided by forward transitional probabilities between adjacent syllables (see Aslin, Saffran, & Newport, 1998). Transitional probabilities were always 1.0 within words and .2 between words.

The two artificial languages differed exclusively in their phonetic structure. In the cv language, the syllables were generated by pairing one plosive (/b d g/) with either one full vowel (/ɑ æ ɛ e u ɔ ʌ/) or one schwa vowel (/ə/). In the vv language, the syllables comprised either one semivowel (/j w/) or one nonlateral approximant (/ð/) as the first segment, paired with either one full or schwa vowel as the second segment. The two languages were symmetrical: For each cv syllable in the former, there was a corresponding vv syllable in the latter. Across the two languages, the consonant and the semivowel had the same place of articulation and were paired with the same full vowel (e.g., /be/ and /dɛ/ in the cv language corresponded to /we/ and /ðɛ/ in the vv language). Two

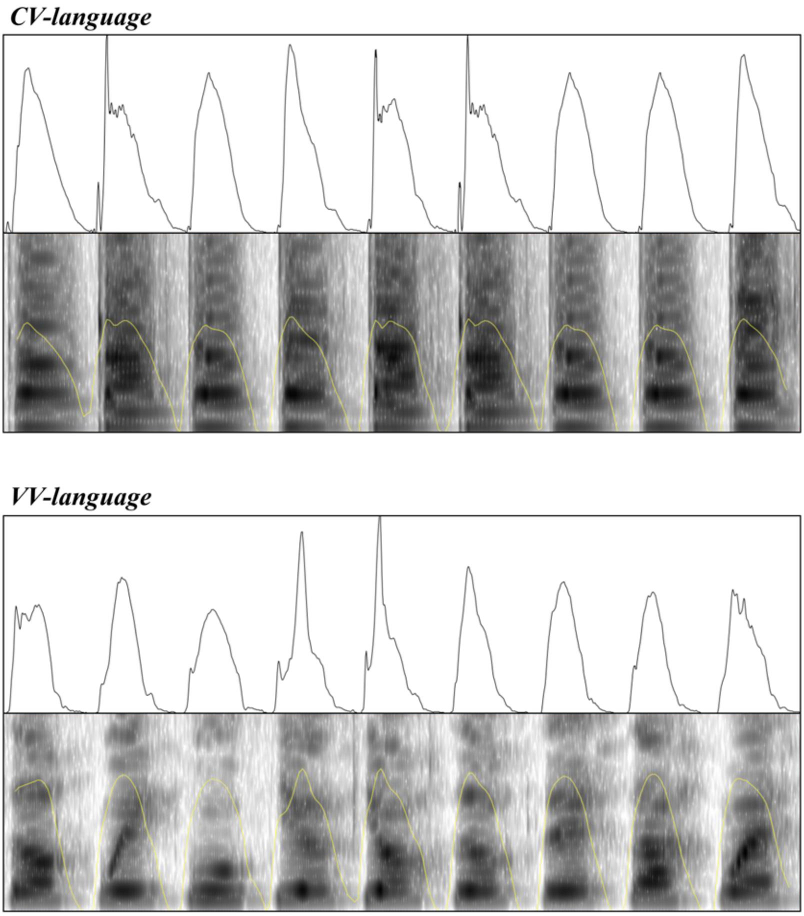

Figure 2 Amplitude envelopes and spectrograms for excerpts in the cv and vv language. [Color figure can be viewed at wileyonlinelibrary.com]

exceptions were made to this rule: (a) /də/ was paired with /ðe/ as /ðə/ would have sounded too similar to the definite article *the* in English and (b) /gæ/ was paired with /jɛ/ because /jæ/ would have been equivalent to *ja*, the Danish word for "yes."

The three plosives were phonetically realized as [b̥ d̥ g̊], the two semivowels were realized as [i̯ u̯], and the nonlateral approximant was realized as the weak noiseless [ð̞], following phonological rules of Danish (Basbøll, 2005). All syllables were recorded individually by a female native speaker of Danish (commonly used speech synthesizers like MBROLA are not available for

Danish), then concatenated in Praat (Boersma & Weenink, 2016). Both the cv and vv syllables had a mean duration of 290 milliseconds ($SD_{cv} = 0.22$, $SD_{vv} = 0.16$). The mean pitch was 236 Hz ($SD = 2.4$) in the cv language and 245.2 Hz ($SD = 2.3$) in the vv language. The amplitude of the individual syllables was root mean square normalized.

Procedure

The experiment consisted of two parts: (a) a training block in which the participants listened to the 11-minute-long speech stream from either the cv language or the vv language (within each native language group, half of the participants were randomly assigned to listen to the former and the other half of them were assigned to listen to the latter) and (b) a test block in which participants were tested on their ability to recognize the six target words of the training block from six nonword foils across 36 two-alternative forced-choice test items. The nonword foils were generated by rearranging syllables from the training words in a new order so that their transitional probabilities were 0. In each of the two-alternative forced-choice test items, the participants were presented acoustically with one word/foil pair. They then had to choose which of the two words most resembled a word from the training speech by clicking with the mouse on the appropriate text box, Word 1 or Word 2, on a computer screen.

To ensure participants' full focus during the training block, we devised a cover task in which the participants were asked to react with a mouse click every time the same syllable occurred twice in a row (e.g., /bebugəgə/). Data from the cover task were not analyzed; the task simply served to focus the participants' attention on the auditory stimuli. The participants sat alone in a quiet room equipped with a computer and listened to the stimuli through headphones during the experiment. The whole procedure was automated and accompanied by step-by-step onscreen instructions that were adapted to Danish, Norwegian, or English, respectively. Prior to the experiment, the participants were asked to fill in a questionnaire about their educational and language background. The three groups did not differ significantly in relation to age, gender, or educational background.

Data Analysis

To quantify the effect of the independent variables on accuracy in the two-alternative forced-choice test, we fit a generalized linear mixed-effects model (which is ideal for binary outcomes) using the glmer function in the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2018). The model had fixed effects for phonetic structure (cv language, vv language),

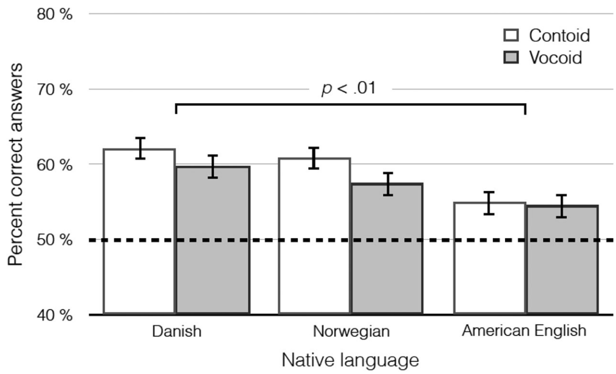

Figure 3 Mean accuracy in the two-alternative forced-choice test by native language and phonetic structure of the artificial language. Error bars indicate standard error of the mean.

native language (Danish, Norwegian, American English), number of syllables (disyllabic, trisyllabic), and an interaction between the first two factors. As random factors, the model had random intercepts for subjects as well as for target words and foils, both nested in phonetic structure. Raw accuracy data were used as the dependent variable.

To get a more detailed picture of our participants' learning of the artificial languages, we also fit a linear mixed-effects model for response times for correct responses in the two-alternative forced-choice test. We started by specifying a maximal model with raw response time data (milliseconds) as our dependent variable and native language, phonetic structure, number of syllables, and the interaction between these three as fixed-effect terms. The random effects remained subjects, target words, and foils as in the previous model. We then performed a stepwise deletion of nonsignificant model terms, using the step function in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) in R.

Results

Figure 3 summarizes the main outcomes from the test phase, and Table 2 reports parameter estimates, confidence intervals, and significance values for all the fixed effects as well as marginal and conditional effect sizes (R^2) for the whole model targeting response accuracy (see Nakagawa & Schielzeth, 2013). On average, participants in all conditions preferred target words to nonword foils in the two-alternative forced-choice test, $M_{\text{target words}} = 0.58$,

Table 2 Summary of fixed-effects coefficient estimates and significance values for outcomes on the two-alternative forced-choice test

Variable	b	SE	CI[a]	z
Intercept	0.23	0.20	[−0.15, 0.62]	1.13
Native language (DA vs. NO)	−0.05	0.13	[−0.37, 0.31]	−0.43
Native language (DA vs. AE)	−0.32	0.13	[−0.55, −0.05]	−2.48*
Phonetic structure (cv vs. vv)	−0.12	0.28	[−0.67, 0.43]	−0.45
Number of syllables (disyllabic vs. trisyllabic)	0.61	0.12	[0.36, 0.92]	4.90***
Native language (NO): Phonetic structure (vv)	−0.04	0.18	[−0.55, 0.28]	−0.23
Native language (AE): Phonetic structure (vv)	0.09	0.18	[−0.31, 0.43]	0.52
Effect size (whole model): $R^2_{\text{marginal}} = .031$, $R^2_{\text{conditional}} = .112$				
Model specifications: Accuracy ~ NativeLanguage * PhoneticStructure + NumberSyllables + (1 | Subject) + (1 | PhoneticStructure:Target) + (1 | PhoneticStructure:Foil).				

Note. DA = Danish; NO = Norwegian; AE = American English; c = contoid (i.e., obstruents, nasals, and lateral approximants); v = vocoid (i.e., full vowels, schwa vowels, semivowels/glides, and nonlateral approximants). [a]CI = Bootstrap confidence interval calculations based on 200 bootstrap replicates [2.5%, 97.5%]. $^{*}p < .05$, $^{***}p < .001$.

$SD = 0.11$, $t(167) = 9.52$, $p < .0001$. Speakers of Danish, Norwegian, and English were accurate in recognizing words from the training language ($M_{\text{Danish}} = 0.60$, $SD = 0.48$; $M_{\text{Norwegian}} = 0.59$, $SD = 0.49$; $M_{\text{English}} = 0.54$, $SD = 0.49$). However, English speakers performed significantly worse than Danish speakers, though still above chance, $t(55) = 3.52$, $p < .001$; there was no significant difference between speakers of Norwegian and Danish. Crucially, we found no significant effect of phonetic structure on performance ($p = .65$), though performance was numerically worse on the vv language across native language groups. A Bayesian factor analysis returning credibility intervals centered around 0 [−0.82, 0.57] and an evidence ratio (Bayes factor) of 29.23 in favor of the null hypothesis suggested that the null result was robust rather than a spurious result due to our experimental design (more details on the Bayesian factor analysis are available in Appendix S2 in the Supporting Information online). Last, the model revealed a highly significant main effect for number of syllables, with generally better performance on trisyllabic words ($M_{\text{disyllabic}} = 0.51$, $SD = 0.49$; $M_{\text{trisyllabic}} = 0.65$, $SD = 0.47$).

Table 3 Likelihood ratio tests and deletion order from the full model for analysis of response times on the two-alternative forced-choice test

Variable	order	*F*	*df*	*p*
Native language	kept	8.77	2, 163.5	< .001
Phonetic structure	5	2.49	1, 162.3	.116
Number of syllables	kept	4.80	1, 3476.5	.028
Native language × phonetic structure	3	1.69	2, 159.6	.186
Native language × number of syllables	2	1.44	2, 3472.7	.235
Phonetic structure × number of syllables	4	2.42	1, 3476	.119
Native language × phonetic structure × number of syllables	1	1.97	2, 3471.7	.138

Effect size (whole model): $R^2_{\text{marginal}} = .019$, $R^2_{\text{conditional}} = .086$

Final model after the stepwise term deletion: ResponseTimes ~ NativeLanguage + NumberSyllables + (1 | Subject). Adding target words and foils as fixed effects rather than random effects to the maximal model did not change the outcome of the stepwise deletion process.

Table 4 Differences of least squares means for significant fixed effects for analysis of response times on the two-alternative forced-choice test

Variable	*b*	*SE*	95% CI	*t*	*df*
Native language					
Danish vs. Norwegian	−194	97.5	[−386.9, −1.9]	−1.99*	160
Danish vs. American English	−411	98.2	[−605.3, −217.5]	−4.19***	165
Norwegian vs. American English	−217	98.5	[−411.6, −22.5]	−2.20*	166
Number of syllables					
disyllabic vs. trisyllabic	112	51.2	[11.8, 212.7]	2.19*	3476

CI = confidence interval. $^{*}p < .05$, $^{***}p < .001$.

Table 3 reports the likelihood ratio tests for fixed effects in the final model along with marginal and conditional effect sizes (R^2) for the whole model targeting response times, and Table 4 reports parameter estimates, confidence intervals, and significant values for differences of least squares means for significant fixed effects. As with the accuracy data, we found a significant effect on the intercept of number of syllables (Table 3), with correct answers to trisyllabic words being given faster than answers to disyllabic words ($M_{\text{disyllabic}}$ = 1,400.80, SD = 1,618.70; $M_{\text{trisyllabic}}$ = 1,300.50, SD = 1,484.10) (Table 4). Moreover, we found a highly significant effect of native language (Table 3). Speakers of Danish were faster than both speakers of Norwegian and of

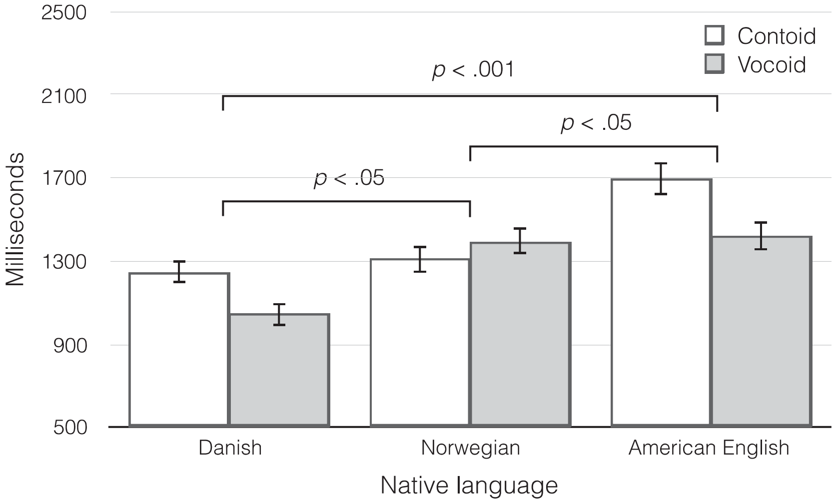

Figure 4 Mean response times in the two-alternative forced-choice test by native language and phonetic structure of the artificial language. Error bars indicate standard error of the mean.

English in choosing the correct word in the test, and speakers of Norwegian were faster than English speakers ($M_{\text{Danish}} = 1{,}192.29$, $SD = 1{,}354.40$; $M_{\text{Norwegian}} = 1{,}427.47$, $SD = 1{,}746.68$; $M_{\text{English}} = 1{,}619.82$, $SD = 1{,}882.16$) (Table 4). We did not find a main effect of phonetic structure nor interactions with other predictors (Table 3). Figure 4 shows the average response times in the various conditions.

Figure 5 shows the average responses to the individual target words and nonword foils. Visual inspection confirmed that accuracy in the lexical discrimination test was generally above chance on trisyllabic words (Words 4 to 6), but not on disyllabic words (Words 1 to 3). Moreover, accuracy was above chance when the target words were presented in combination with disyllabic nonword foils but not when the foils were trisyllabic (with the exception of Foil 6).

General Discussion

In this study, we explored the hypothesis that highly vocalic speech may pose a challenge for speech segmentation via statistical learning. This study was primarily motivated by research suggesting that Danish, which has an unusual vocalic phonetic structure, may be intrinsically difficult to acquire for both first language learners (e.g., Bleses et al., 2008; Bleses, Basbøll, Lum, et al., 2011) and second language learners (e.g., Gooskens et al., 2010; Grønnum,

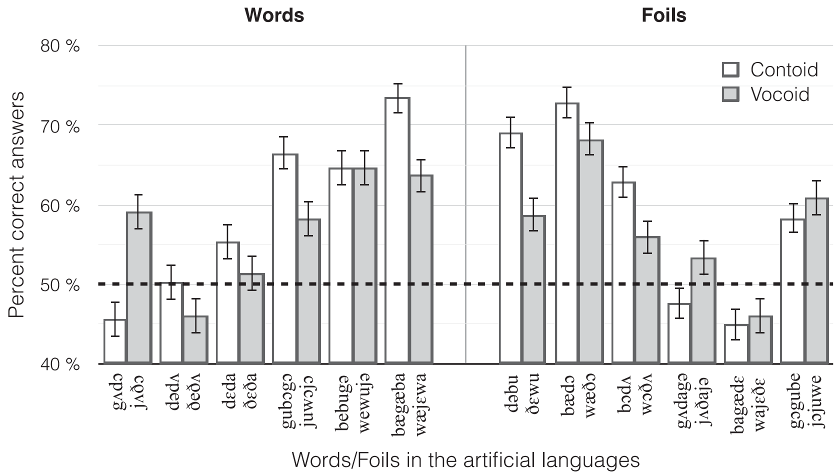

Figure 5 Accuracy on the two-alternative forced-choice test for individual words and foils by phonetic structure of the target language. Error bars indicate standard error of the mean.

2003). Literature suggesting that contoids are more useful for segmentation than vocoids (e.g., Bonatti et al., 2005; Nespor et al., 2003) raised the question of whether the reduced occurrence of contoids in Danish may render the speech input intrinsically less segmentable. To investigate this hypothesis, we tested adult speakers of Danish, Norwegian, and English on their ability to use statistical learning to segment two artificial speech streams consisting of either contoid–vocoid or vocoid–vocoid syllables. Our crosslinguistic experiment was preceded by corpus analyses that showed that learners of Danish are presented with the challenge of positing syllable and word boundaries in vocalic sequences more frequently than are English learners.

The main finding of our experiment was that the ratio of contoids to vocoids in the two languages did not affect the participants' performance in distinguishing words from nonwords. The vv language was learned as accurately as the cv language. Therefore, we found no support for the contoid-biased segmentation hypothesis, that is, the idea that a high ratio of vocoids versus contoids reduces the amount of information available for statistical learning. The relatively large sample size compared to other studies of this kind, together with the ability of our design to detect significant differences both in the performance across native language groups and across number of syllables, suggests that the lack of an effect of phonetic structure (cv vs. vv) is not spurious; rather, it provides clear evidence for the lack of a difference in inherent segmentability between the two artificial languages. Implications of this finding for a weaker version of our hypothesis, the syllable diffusion hypothesis, together with alternative interpretations of the results are discussed below.

Three minor findings also emerged from the experiment. First, we found a significant effect for the participants' native language on both accuracy (Danish and Norwegian speakers scored higher than English speakers on the two-alternative forced-choice test) and response times (Danish speakers were faster than Norwegian speakers, and both groups were faster than English speakers). This is in line with our prediction that familiarity with phonetic–phonological properties of the artificial languages would be an advantage for our Danish-speaking participants and, to a lesser degree, for our Norwegian-speaking participants due to the affinity between the two languages (cf. Vroomen, Tuomainen, & de Gelder, 1998). Speakers of Norwegian are also more likely than speakers of English to be exposed to Danish.

Second, we found that performance was consistently better on trisyllabic words compared to disyllabic ones. The presence of more statistical information in the trisyllabic words may have helped our participants glue the words together to a greater extent compared to the disyllables (cf. Johnson & Tyler, 2010; Frank, Tily, Arnon, & Goldwater, 2010; Lew-Williams & Saffran, 2012, for a discussion of transitional probabilities and word length). Third, we found no evidence for the idea that the vocoid bias observed in Danish 20-month-olds (Højen & Nazzi, 2016), which is possibly evidence of the ambient language shaping the child's phonological system in early childhood, made the Danish participants better equipped for segmenting the vv language than the cv language.

The lack of evidence in favor of the consonant-biased segmentation hypothesis may provide initial support for the syllable diffusion hypothesis, that is, the idea that sequences of vocoids, by being more prone to coarticulation, assimilation, and reduction, may make the statistical structure of language particularly susceptible to degradation. We suggest that sequences of vocoids in Danish may make crosssyllable transitions ambiguous because of two processes: (a) changes in the sonority profile of the syllable brought about by the weakening of contoids to vocoids, with consequently increased chances of coarticulation and assimilation, and (b) the loss of syllables due to the frequent deletion of schwa vowels.

First, the weakening of contoids into vocoids generally affects the sonority profile of the speech stream (e.g., Basbøll, 2012; Kjærbæk, Boeg Thomsen, Lambertsen, & Basbøll, 2015). Each segment in speech has a specific degree of

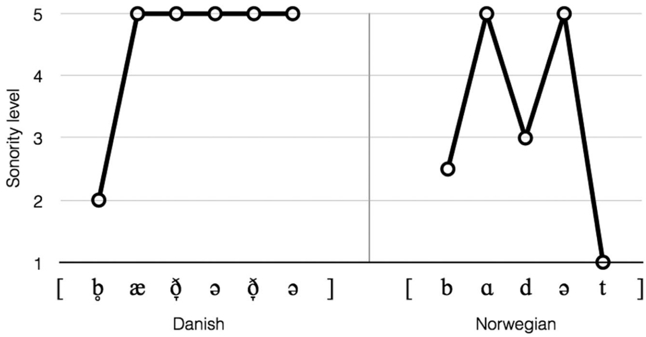

Figure 6 Sonority profile of the Danish verb *badede* "(I) bathed" and its Norwegian cognate *badet* following Basbøll's (2005, 2012) sonority syllable model. The five sonority levels are: (a) +SG, (b) −SG −voi, (c) +voi −son, (d) +son −voc, and (e) +voc. Norwegian /b/ lies between Levels 2 and 3 because it is half-voiced (Kristoffersen, 2000). [b̥æːð̞əð̞ə] is a very distinct pronunciation of the Danish past tense, and in casual speech the schwa vowels are typically assimilated to the preceding segments, resulting in [b̥æːð̞ð̩] or even [b̥æːð̩].

intrinsic sonority in relation to other segments. Contoids such as obstruents are the least sonorous segments in speech whereas vocoids are the most sonorous. Prototypically, segments with maximal sonority constitute the peak of a syllable while those with minimal sonority constitute the nonpeak. The alternation of nonpeaks and peaks in speech, which is prototypically realized in cv syllables, maximizes the difference in salience between the two classes of segments in syllables, making the syllabic structure robust, that is, more resilient to degradation, and therefore to loss of information, compared to sequences in which this alternation is lost (e.g., Wright, 2004). However, segments that are affected by weakening processes lose their consonantal nature and gain in sonority. When this happens, the distance in perceptual salience between segments is considerably reduced, and the general sonority profile of speech becomes higher and more uniform.

This effect on the sonority profile of Danish speech can be easily illustrated by means of the sonority syllable model (Basbøll, 2005, 2012; for other foundational work on sonority, see also Clements, 1990; Ohala, 1992; Parker, 2002) presented in Figure 6. In the model, speech segments are organized hierarchically according to their intrinsic sonority level, from the least to the most sonorous, as follows:

1. voiceless obstruents (nonsonorants) with widely spread glottis [+SG],

2. voiceless obstruents without widely spread glottis [−SG −voi],
3. voiced obstruents [+voi −son],
4. nonvocoid sonorants [+son −voc], and
5. vocoids [+voc] (the preceding levels include contoids).

Increases in sonority due to consonant weakening result in a segment moving up the scale. This effect can be better appreciated by comparing two cognate words: Danish *badede* [b̥æːð̞əð̞ə] and Norwegian *badet* [bɑːdət], "I bathed." It is evident that the Norwegian word *badet* has a clear structure marked by the contoid–vocoid alternation in sonority. In Danish, however, everything after the initial obstruent is realized as a vocoid. When the contoid–vocoid alternation is lost, the syllabic structure of the word becomes less robust and less transparent, with obvious consequences for statistical learning.

The degradation of the syllabic structure may be aggravated by the pervasive deletion that some vocoids are susceptible to in Danish speech. In particular, the schwa vowel /ə/, which occurs in unstressed syllables in Danish, is ubiquitously assimilated to neighboring sonorants, causing these to become syllabic, for example, *badede* [b̥æːð̞əð̞ə]. The two schwas are normally assimilated to the two preceding [ð̞]s, which overtake the syllabic role of the entire third syllable. The result is the highly reduced form [b̥æːð̞ð̞], or even [b̥æːð̞], in some instances of casual speech. Crucially, this assimilation is not explicitly marked from a phonetic perspective, except perhaps for a slight lengthening of the [ð̞] sound, thus possibly tricking the beginner listener (child or adult) into believing that the word is monosyllabic or disyllabic rather than trisyllabic (Grønnum, 2003). Norwegian nouns and verbs, on the other hand, are not so prone to such radical reductions (cf. Kristoffersen, 2000). The (phonetic) loss of segment/syllables thus can result in substantial changes to the statistical structure of language, for instance, by affecting the segment count of words in speech (Cutler, 1998; Johnson, 2004), thereby reducing the information available for statistical learning based on computations of transitional probabilities over syllables. The fact that participants' performance was consistently better on trisyllabic words relative to disyllabic words in our study is in line with this idea, suggesting that sensitivity to words in continuous speech via transitional probabilities is higher when more statistical information is present in the speech stream.

The method used in our study was not sensitive to the processes described in our syllable diffusion hypothesis: The syllabic structure of our vv language was indeed robust and predictable and did not allow for the reduction processes described above. Having rejected the contoid-biased segmentation hypothesis, we suggest that additional studies are now needed to illuminate issues related to its weaker version, and we discuss possible designs for additional experiments below.

Alternative Interpretations

Some alternative interpretations of our findings must also be considered. First, it is important to note that our design did not allow us to test a more temporal–acoustic aspect of our contoid-biased segmentation hypothesis, namely, that processing highly vocalic speech may be harder simply because it is harder to identify temporally precise boundaries between adjacent vocoids. More specifically, coarticulations and assimilations of neighboring vocoids may reduce segmentability not because they affect the statistical structure of language as suggested by our syllable diffusion hypothesis but simply because of the lack of clear temporal–acoustic cues to segment boundaries between vocoids. However, small dips in amplitude between the concatenated syllables in our artificial languages (see Figure 2) did not allow us to test this hypothesis explicitly.

Another alternative interpretation of our results concerns the fact that our participants may have perceived and processed the semivowels (i.e., /j w ð/) in our vv syllables as phonological consonants (i.e., nonpeaks) rather than as phonetic vowels as we originally intended them to do. Examples of semivowel–vowel monosyllables, in which semivowels serve as phonological consonants, are not rare in English (e.g., the word *we*) or in Danish and Norwegian (e.g., the word *ja*, "yes"). Familiarity with these constructions from their native languages may have facilitated the segmentation process for our participants, thus making the vv language as segmentable as the cv language. Our motivation for using a mix of semivowels, full vowels, and schwa vowels (rather than full vowels alone) in the vv syllables stemmed from the fact that a speech stream consisting only of full vowels would not have been representative of Danish or of any other natural language. Moreover, adjacent full vowels would have likely been perceived as diphthongs. We also excluded the possibility of testing our participants on VC sequences, as those are likely to be perceived as CV streams in fluent speech (cf. Bonatti et al., 2005). Future studies using variations on the current experimental design (e.g., using natural language stimuli with infants) may help reduce the impact of phonological biases of this kind.

In general, our results seem to be inconsistent with the studies reviewed in the introductory section, which have shown that adults correctly segment words with cv syllabic structure but not words with vv syllabic structure from continuous speech (Bonatti et al., 2005; Mehler et al., 2006; Toro, Nespor, Mehler, & Bonatti, 2008; see also Pons & Toro, 2010; Toro, Shukla, Nespor, & Endress, 2008). This is in spite of the fact that the paradigm used in our study was similar to the ones used in the aforementioned investigations. However, these studies were designed specifically to test how statistics would be computed over segmental tiers (i.e., over either C_i or V_i in $C_1V_1C_2V_2C_3V_3$-sequences) rather than over adjacent syllable tiers, as in our case (i.e., XV_1 XV_2 XV_3, where X is either a contoid or a vocoid across the two artificial languages). That is, prior studies were designed to tap into questions at a higher level of representation: for instance, the phonological/functional role of consonants and vowels in speech (whether they form peaks or nonpeaks of syllables) or even questions concerning the sensitivity to the similarity of elements in speech (Newport & Aslin, 2004).

Limitations and Future Research

Our study was designed to test the contoid-biased segmentation hypothesis about the difficulty associated with segmenting highly vocalic speech. The contoid-biased segmentation hypothesis pertains to characteristics of the individual segments in speech: We tested this hypothesis in our study by manipulating the nature of segments across the two artificial languages. To test weaker versions of our hypothesis, including our syllable diffusion hypothesis, future work should focus on modeling more closely the disadvantages presented by vocalic sequences, for instance, in Danish speech, through artificial languages. This would presumably involve moving away from the stimuli used in traditional segmentation paradigms (e.g., Saffran et al., 1996) and using artificial speech streams that (a) have less prominent cues to syllable boundaries and a less robust syllabic structure; (b) allow for changes in the statistical properties of speech, for instance, by being susceptible to the elision of unstressed vocoids or of entire syllables; and (c) can simulate degraded segmentability of the speech stream either through the addition of noise (e.g., white/brown noise) or through (simulated) lenition processes.[4] Our data also showed no sign of a Danish vocalic advantage over the Norwegian and English groups, suggesting that the bias for vocoids observed in Danish-learning children is, if at all persisting into adulthood, not reflected in this kind of task. The use of artificial speech stimuli that model characteristics of Danish more closely may also make it possible to address this issue more explicitly.

Another priority for future research concerns the fact that prosodic elements such as *stød*, a form of creaky voice phonation occurring in Danish in some long vowels and in some sonorants following short vowels (Basbøll, 2005), may function as a segmentation cue in otherwise ambiguous crossword transitions, thus making long vocalic stretches easier to segment. A recent study by Gómez, Mok, Ordin, Mehler, and Nespor (2018) found that speakers of Cantonese (a tonal language) rely more on cues carried by vocoids than contoids when segmenting speech, especially when these vocoids are tone bearing. Therefore, it would be important to investigate whether Danish *stød* (which was intentionally not present in the stimuli used in this study) could provide vocoids with better segmentation cues in the same way that tone-bearing vocoids aid segmentation in tonal languages like Cantonese.

Furthermore, it is important to mention that other methodological aspects of our study may have affected our results. For instance, having longer familiarization times or fewer/more target words in the test phase than in other comparable studies may have played a role in minimizing the differences in performance in the two conditions (e.g., Saffran et al., 1996, used a 21-minute-long training phase followed by as many as 72 test items). However, there is also evidence that long exposure times can lead to worse segmentation accuracy (e.g., Endress & Bonatti, 2007). Moreover, the use of words of different length may have made the task generally harder (Hoch, Tyler, & Tillman, 2013), and the use of a cover task in our training block may have defeated our purpose of focusing the listeners' attention on the task and accidentally affected their segmentation performance in a negative way. For example, Toro, Sinnett, and Soto-Faraco (2011) found that segmentation performance got significantly worse when their participants' attention was diverted from the speech stream during training. The lack of a gold standard in the field makes this kind of procedures more susceptible to methodological biases. Further studies will have to focus on controlling for these possible confounding factors. Last, the three native languages included in our study all have relatively high ratios of vocoids to contoids in speech. Future studies may thus have to include languages with markedly larger ratios of contoids to vocoids (e.g., Serbo-Croatian, Galician).

Conclusion

In conclusion, our study did not provide support for the hypothesis that vv syllables are inherently harder to segment than cv syllables in a statistical learning task. Given the lack of evidence in support of the contoid-biased segmentation hypothesis, future research is needed to focus on these less direct ways in which vocoid sequences can affect speech segmentation, namely, our syllable diffusion hypothesis. The broader implications of these findings are not only important for the study of Danish but also for bringing into focus the role of crosslinguistic variability in the study of speech segmentation in both first and second language learners. Our results seem to suggest that statistical learning is robust and reliable independent of the phonetic properties of the input language as long as its statistical properties are informative. At the same time, the structural properties of speech can vary considerably across natural languages in terms of their relative informativeness and processability (Bates & MacWhinney, 1989; Cutler & Carter, 1987). Cues that are more available and reliable should result, all else being equal, in processing and acquisition advantages. At the same time, reduced cue salience may pose a challenge for processing and learning (Bates et al., 2001). Investigating how domain-general learning mechanisms cope with structural differences across languages is therefore essential to the study of language learning.

Final revised version accepted 11 July 2018

Notes

1 Furthermore, Grønnum adds one abstract schwa phoneme, but there are, at a concrete phonological level, two schwa vowels, namely, /ə/ and /ɐ/ (Basbøll, 2005).
2 The vocoid-to-contoid ratio for Norwegian was not reported by Bleses, Basbøll, Lum, et al. (2011); we therefore computed the ratio ourselves using the same procedure.
3 Contrary to what many previous studies of word segmentation via statistical learning have done, we varied the length of our nonsense words in a way that better approximated real-world language learning situations. We were particularly interested in testing the effect of word length on statistical learning performance because syllabic reduction is pervasive in casual Danish speech. This point is addressed at length in the discussion section.
4 We thank an anonymous reviewer for these suggestions.

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324. https://doi.org/10.1111/1467-9280.00063

Basbøll, H. (2005). *The phonology of Danish*. Oxford, UK: Oxford University Press.

Basbøll, H. (2012). Monosyllables and prosody: The sonority syllable model meets the word. In T. Stolz, N. Nau, & C. Stroh (Eds.), *Studia typologica. Monosyllables: From phonology to typology* (pp. 13–41). Berlin, Germany: De Gruyter.

Basbøll, H., Bleses, D., Cadierno, T., Jensen, A. J., Ladegaard, H. J., Madsen, T. O., . . . Thomsen, P. (2002). The Odense language acquisition project. *Child Language Bulletin*, *22*, 11–12.

Bates, E., Devescovi, A., & Wulfeck, B. (2001). Psycholinguistics: A cross-language perspective. *Annual Review of Psychology*, *52*, 369–396. https://doi.org/10.1146/annurev.psych.52.1.369

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–76). New York, NY: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bauer, L. (1983). Consonant strength hierarchies and Danish. *Nordic Journal of Linguistics*, *6*, 115–128. https://doi.org/10.1017/S0332586500001037

Bleses, D., & Basbøll, H. (2004). The Danish sound structure: Implications for language acquisition in normal and hearing impaired populations. In E. Schmidt, U. Mikkelsen, I. Post, J. B. Simonsen, & K. Fruensgaard (Eds.), *Brain, hearing and learning. 20th Danavox Symposium* (pp. 165–190). Copenhagen, Denmark: Holmen Center Tryk.

Bleses, D., Basbøll, H., Lum, J., & Vach, W. (2011). Phonology and lexicon in a cross-linguistic perspective: the importance of phonetics: A commentary on Stoel-Gammon's "Relationships between lexical and phonological development in young children." *Journal of Child Language*, *38*, 61–68. https://doi.org/10.1017/S0305000910000437

Bleses, D., Basbøll, H., & Vach, W. (2011). Is Danish difficult to acquire? Evidence from Nordic past-tense studies. *Language and Cognitive Processes*, *26*, 1193–1231. https://doi.org/10.1080/01690965.2010.515107

Bleses, D., & Trecca, F. (2016). Early acquisition of Danish in a cross-Scandinavian perspective: A psycholinguistic challenge? In H.-O. Enger, M. I. N. Knoph, K. E. Kristoffersen, & M. Lind (Eds.), *Helt fabelaktig! Festskrift til Hanne Gram Simonsen på 70-årsdagen* (pp. 13–28). Oslo, Norway: Novus Forlag.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T., & Basbøll, H. (2008). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language*, *35*, 619–650. https://doi.org/10.1017/S0305000908008714

Bleses, D., Vach, W., Wehberg, S., Faber, K., & Madsen, T. O. (2007). *Tidlig kommunikativ udvikling: Værktøj til beskrivelse af sprogtilegnelse [Early communicative development: A tool for assessing language development]*. Odense, Denmark: University Press of Southern Denmark.

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer (Version 6.0.14) [Computer program]. Retrieved from http://www.praat.org

Bonatti, L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, *16*, 451–459. https://doi.org/10.1111/j.0956-7976.2005.01556.x

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (2015). Call me Alix, not Elix: Vowels are more important than consonants in own-name recognition at 5 months. *Developmental Science*, *18*, 587–598. https://doi.org/10.1111/desc.12242

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, *61*, 93–125. https://doi.org/10.1016/S0010-0277(96)00719-6

Carroll, S. E. (2004). Segmentation: Learning how to “hear words” in the L2 speech stream. *Transactions of the Philological Society*, *102*, 227–254. https://doi.org/10.1111/j.0079-1636.2004.00136.x

Clements, G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology 1: Between the grammar and physics of speech* (pp. 283–333). Cambridge, UK: Cambridge University Press.

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*, 233–262. https://doi.org/10.1016/j.cognition.2004.08.005

Cutler, A. (1998). The recognition of spoken words with variable representation. In D. Duez (Ed.), *Proceedings of the ESCA workshop on sound patterns of spontaneous speech* (pp. 83–92). Aix-en-Provence, France: Université de Aix-en-Provence.

Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142. https://doi.org/10.1016/0885-2308(87)90004-0

Demetras, M. (1989). *Working parents’ conversational responses to their two-year-old sons* (Working paper). Unpublished manuscript, Department of Speech and Hearing Sciences, University of Arizona, Tucson, AZ.

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, *105*, 247–299. https://doi.org/10.1016/j.cognition.2006.09.010

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108. https://doi.org/10.1016/j.dr.2015.05.002

Frank, M. C., Tily, H., Arnon, I., & Goldwater, S. (2010). Beyond transitional probabilities: Human learners apply a parsimony bias in statistical word segmentation. *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, *32*, 760–765.

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, *61*, 191–218. https://doi.org/10.1146/annurev.psych.093008.100408

Gómez, D. M., Mok, P., Ordin, M., Mehler, J., & Nespor, M. (2018). Statistical speech segmentation in tone languages: The role of lexical tones. *Language and Speech*, *61*, 84–96. https://doi.org/10.1177/0023830917706529

Gonzalez-Gomez, N., Hayashi, A., Tsuji, S., Mazuka, R., & Nazzi, T. (2014). The role of the input on the development of the LC bias: A crosslinguistic comparison. *Cognition*, *132*, 301–311. https://doi.org/10.1016/j.cognition.2014.04.004

Gooskens, C., van Heuven, V. J., van Bezooijen, R., & Pacilly, J. J. A. (2010). Is spoken Danish less intelligible than Swedish? *Speech Communication*, *52*, 1022–1037. https://doi.org/10.1016/j.specom.2010.06.005

Grønnum, N. (1998). Danish: Illustrations of the IPA. *Journal of the International Phonetic Association*, *28*, 99–105. https://doi.org/10.1017/S0025100300006290

Grønnum, N. (2003). Why are the Danes so hard to understand? In H. G. Jacobsen, D. Bleses, T. O. Madsen, & P. Thomsen (Eds.), *Take Danish—for instance* (pp. 119–130). Odense, Denmark: University Press of Southern Denmark.

Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy*, *14*, 439–456. https://doi.org/10.1080/15250000902996532

Hilton, N. H., Schüppert, A., & Gooskens, C. (2011). Syllable reduction and articulation rates in Danish, Norwegian and Swedish. *Nordic Journal of Linguistics*, *34*, 215–237. https://doi.org/10.1017/S0332586511000175

Hoch, L., Tyler, M. D., & Tillmann, B. (2013). Regularity of unit length boosts statistical learning in verbal and nonverbal artificial languages. *Psychonomic Bulletin & Review*, *20*, 142–147. https://doi.org/10.3758/s13423-012-0309-8

Hochmann, J. -R., Benavides-Varela, S., Fló, A., Nespor, M., & Mehler, J. (2017). Bias for vocalic over consonantal information in 6-month-olds. *Infancy*, *23*, 136–151. https://doi.org/10.1111/infa.12203

Hochmann, J. -R., Benavides-Varela, S., Nespor, N., & Mehler, J. (2011). Consonants and vowels: Different roles in early language acquisition. *Developmental Science*, *14*, 1445–1458. https://doi.org/10.1111/j.1467-7687.2011.01089.x

Højen, A., & Nazzi, T. (2016). Vowel bias in Danish word-learning: Processing biases are language-specific. *Developmental Science*, *19*, 41–49. https://doi.org/10.1111/desc.12286

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous speech: Data and analysis. Proceedings of the 1st session of the 10th international symposium* (pp. 29–54). Tokyo, Japan: The National International Institute for Japanese Language.

Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, *13*, 339–345. https://doi.org/10.1111/j.1467-7687.2009.00886.x

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences*, *3*, 323–328. https://doi.org/10.1016/S1364-6613(99)01363-7

Keidel, J. L., Jenison, R. L., Kluender, K. R., & Seidenberg, M. S. (2007). Does grammar constrain statistical learning? Commentary on Bonatti, Peña, Nespor, and Mehler (2005). *Psychological Science*, *18*, 922–923. https://doi.org/10.1111/j.1467-9280.2007.02001.x

Kim, Y. J., & Sundara, M. (2015). Segmentation of vowel-initial words is facilitated by function words. *Journal of Child Language*, *42*, 709–733. https://doi.org/10.1017/S0305000914000269

Kjærbæk, L., Boeg Thomsen, D., Lambertsen, C., & Basbøll, H. (2015). Sonority and early words: The sonority syllable model applied to an acquisitional project with Danish children. *Italian Journal of Linguistics*, *27*, 61–84.

Kjærbæk, L., Christensen, R. D., & Basbøll, H. (2014). Sound structure and input frequency impact on noun plural acquisition: Hypotheses tested on Danish children across different data types. *Nordic Journal of Linguistics*, *37*, 47–86. https://doi.org/10.1017/S0332586514000067

Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford, UK: Oxford University Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. https://www.jstatsoft.org/article/view/v082i13

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition*, *122*, 241–246. https://doi.org/10.1016/j.cognition.2011.10.007

Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358–368.

MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science*, *288*, 527–531. https://doi.org/10.1126/science.288.5465.527

MacWhinney, B. (2000). *The CHILDES project* (3rd ed.). Mahwah, NJ: Erlbaum.

MacWhinney, B., & Bates, E. (Eds). (1989). *The crosslinguistic study of sentence processing*. New York, NY: Cambridge University Press.

MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of Child Language*, *12*, 271–295.

Mattys, S. L., & Juszcyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology*, *27*, 644–655. https://doi.org/10.1037/0096-1523.27.3.644

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. L. (2006). The “soul” of language does not use statistics: Reflections on vowels and consonants. *Cortex*, *42*, 846–854. https://doi.org/10.1016/S0010-9452(08)70427-1

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nazzi, T., Dilley, L. C., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants’ segmentation of verbs from fluent speech. *Language and Speech*, *48*, 279–298. https://doi.org/10.1177/00238309050480030201

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *2*, 221–247.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Nishibayashi, L.-L., & Nazzi, T. (2016). Vowels, then consonants: Early bias switch in recognizing segmented word forms. *Cognition*, *155*, 188–203. https://doi.org/10.1016/j.cognition.2016.07.003

Ohala, J. J. (1992). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In M. Ziolkowsky, M. Noske, & K. Deaton (Eds.), *Papers from the Chicago Linguistic Society* (pp. 319–339). Chicago, IL: Chicago Linguistic Society.

Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Erlbaum.

Parker, S. G. (2002). *Quantifying the sonority hierarchy* (Unpublished doctoral dissertation). University of Massachusetts, Amherst, MA.

Pike, K. (1943). *Phonetics. University of Michigan Publication in Language and Literature* (Vol. 21). Ann Arbor, MI: University of Michigan Press.

Plunkett, K. (1985). *Preliminary approaches to language development*. Århus, Denmark: Århus University Press.

Plunkett, K. (1986). Learning strategies in two Danish children's language development. *Scandinavian Journal of Psychology*, *27*, 64–73 https://doi.org/10.1111/j.1467-9450.1986.tb01188.x

Pons, F., & Toro, J. M. (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition*, *116*, 361–367. https://doi.org/10.1016/j.cognition.2010.05.013

R Core Team (2018). R: A language and environment for statistical computing (Version 3.5.0) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Ragnarsdóttir, H., Simonsen, H. G., & Plunkett, K. (1999). The acquisition of past tense morphology in Icelandic and Norwegian children: An experimental study. *Journal of Child Language*, *26*, 577–618.

Rischel, J. (1970/2009). Consonant gradation: A problem in Danish phonology and morphology. In N. Grønnum, F. Gregersen, & H. Basbøll (Eds.), *Sound structure in language* (pp. 26–43). Oxford, UK: Oxford University Press.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 906–914. https://doi.org/10.1002/wcs.78

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621. https://doi.org/10.1006/jmla.1996.0032

Schüppert, A., Hilton, N. H., & Gooskens, C. (2016). Why is Danish so difficult to understand for fellow Scandinavians? *Speech Communication*, *79*, 47–60. https://doi.org/10.1016/j.specom.2016.02.001
Seidl, A., & Johnson, E. (2008). Boundary alignment enables 11-month-olds to segment vowel initial words from speech. *Journal of Child Language*, *35*, 1–24. https://doi.org/10.1017/S0305000907008215
Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J. L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, *35*, 869–902. https://doi.org/10.1017/S0305000908008763
Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.
Stoel-Gammon, C. (1998). Sounds and words in early language acquisition. The relationship between lexical and phonological development. In P. Rhea, S. F. Warren, & M. E. Fey (Eds.), *Exploring the speech-language connection* (pp. 25–52). Baltimore, MD: Paul H. Brookes.
Stoel-Gammon, C. (2011). Relationship between lexical and phonological development in young children. *Journal of Child Language*, *28*, 1–34. https://doi.org/10.1017/S0305000910000425
Swingley, D. (1999). Conditional probability and word discovery: A corpus analysis of speech to infants. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st Annual Meeting of the Cognitive Science Society* (pp. 724–729). Mahwah, NJ: Erlbaum.
Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, *39*, 706–716. https://doi.org//10.1037/0012-1649.39.4.706
Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*, *19*, 137–144. https://doi.org/10.1111/j.1467-9280.2008.02059.x
Toro, J. M., Shukla, M., Nespor, M., & Endress, A. D. (2008). The quest for generalizations over consonants: Asymmetries between consonants and vowels are not the by-product of acoustic differences. *Perception & Psychophysics*, *70*, 1515–1525. https://doi.org/10.3758/PP.70.8.1515
Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2011). Generalizing linguistic structures under high attention demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 493–501. https://doi.org/10.3758/PP.70.8.1515
Trecca, F., Bleses, D., Højen, A., Madsen, T. O., & Christiansen, M. H. (2018). *When too many vowels impede language processing: An eye-tracking study of Danish-learning children*. Manuscript submitted for publication.
Trecca, F., Bleses, D., Madsen, T. O., & Christiansen, M. H. (2018). Does sound structure affect word learning? An eye-tracking study of Danish learning toddlers. *Journal of Experimental Child Psychology*, *167*, 180–203. https://doi.org/10.1016/j.jecp.2017.10.011

van Ooijen, B. (1994). *The processing of vowels and consonants*. The Hague, Netherlands: Holland Academic Graphics.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, *38*, 133–149. https://doi.org/10.1006/jmla.1997.2548

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. M. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 34–57). New York, NY: Cambridge University Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1. Rationale for the Contoid–Vocoid Distinction.
Appendix S2. Bayesian Factor Analysis.

Appendix: Accessible Summary (also publicly available at https://oasis-database.org)

Are Languages With Many Vowel Sounds Particularly Hard to Learn?

What This Research Was About And Why It Is Important

The Danish language is not typical of many other languages because it has an unusually large number of vowel sounds (such as *a*, *e*, or *o*) relative to consonants (such as *b*, *d*, or *k*). The large proportion of vowels has been claimed to make it hard for first and second language learners to “extract” and subsequently learn words from spoken Danish. In this study, the researchers tested this idea by investigating whether spoken words composed of alternating consonants and vowels (e.g., *bubogo*, *begeba*) were easier to recognize than words composed of only vowels (e.g., *iuuouo*, *ieueua*). The researchers predicted that learners in the consonant and vowel group would learn to recognize more words than those in the vowel-only group. However, the researchers found that the number of vowels did not make a difference.

What the Researchers Did

- The researchers tested three groups of adults (university students) to compare how speakers of languages that are more or less similar to Danish learned to recognize spoken Danish words:

  - 56 speakers of Danish (who provided a performance level for those who speak Danish natively),
  - 56 speakers of Norwegian, a language closely related to Danish but with fewer vowels than Danish, and
  - 56 speakers of English, a language less related to Danish than Norwegian and also with fewer vowels than in Danish.
- The researchers presented participants with two tasks within the same session.
- Participants first listened to an 11-minute recording of many repetitions of 36 nonsense words (i.e., made up words sounding like real Danish words). The words occurred in a random order, with no pauses between words.
  - For half of the participants in each group, the nonsense words consisted of both consonants and vowels, in alternation (as in *begeba*), typical of Norwegian and English.
  - For the other half, the nonsense words consisted only of vowels (as in *ieueua*), typical of Danish.
- Immediately after, participants heard six pairs of nonsense words; in each pair, one word had occurred in the speech stream while the other was similar but had not occurred previously.
  - Participants selected which word they heard previously, and the researchers analyzed how accurately and quickly participants recognized the previously heard words.

What the Researchers Found

- Regardless of their native language (Danish, Norwegian, or English), all participants were equally good at recognizing the nonsense words experienced both in the consonant and vowel speech stream and in the vowel-only stream (54–60% correct). Thus, the two streams of words were equally easy to learn, despite one of them having a large number of vowel sounds (typical of Danish).
- Language distance played a role too: Speakers of Norwegian (whose performance generally did not differ from that by Danish speakers) were more accurate and also faster to recognize previously heard words than English speakers.

Things to Consider

- The two streams of nonsense words had a very repetitive structure that is quite far from that of actual Danish; this perhaps made words easier to store in memory than in a real language.
- If a real language had *only* vowels, this would dramatically reduce the information available to a learner, as many vowels can result in indistinct or unreliable pronunciation. For example, it is harder to pronounce the vowel-only English word *aiouea* (a rare tree) compared to a word like *aerial*, despite the words having the same number of letters.
- While many vowels did not make individual words less recognizable over a short time, having many vowels may make languages less "information-rich" overall and harder to learn.

How to cite this summary: Trecca, F., & Christiansen, M. H. (2018). Are languages with many vowel sounds particularly hard to learn? *OASIS Summary* of Trecca et al. in *Language Learning*. https://oasis-database.org

This summary has a CC BY–NC–SA license.