CREAKY VOICE IN L2 ENGLISH AND L1 DUTCH

Koen Sebregts, Hielke Vriesendorp, Hugo Quené, Yosiane White

ILS, Utrecht University
k.sebregts@uu.nl, h.a.d.vriesendorp@uu.nl, h.quene@uu.nl, y.z.white@uu.nl

ABSTRACT

While creaky voice is a well-known stylistic-prosodic feature of American English, its use in other languages is under-researched. In Dutch, it has been claimed to be rare and idiosyncratic. Meanwhile, in L2 acquisition studies, creak has been shown to exhibit L1-L2 transfer.

We investigate the prevalence of creaky voice and its development over time in young female speakers of L1 Dutch and L2 English, who had previously been shown to converge on L1/L2 segmental features longitudinally.

Automatic detection of f0 established individual distributions of creaky and modal phonation. Linear mixed-effects models of relative creak prevalence showed that effects of language (L1 vs L2) and style (read vs spontaneous speech) were limited. Rather, speakers showed similarly high levels of creak in their two languages, and stability over time. The results suggest creak is idiosyncratic, rather than gradually acquired or converged on, but also that it is far from rare in Dutch.

Keywords: creaky voice, phonation, L2 acquisition, Dutch, English

1. INTRODUCTION

1.1 Background

Creak and creaky voice are now widely assumed to be umbrella terms for a number of types of non-modal voice qualities [1-3]. Most of these share elements such as low fundamental frequency and some irregularity in glottal pulsing; despite acoustic differences, they appear to be a perceptually coherent category [4].

Cross-linguistically, creak and creaky voice can have anything from phonemic to discourse or affective functions [5]. In English, it is well-known to have prosodic and pragmatic functions, as a boundary marker [6] and turn-taking signal [7], as well as stylistic and sociolinguistic ones. Available as an identity marker of upper-middle class RP speakers to Chicana/o gang members alike [8, 9], it is popularly most associated with young female speech in American English and has received ample attention as a result, both outside of and within linguistic research [10-12].

Considerably less is known about the use of creaky voice in other languages, particularly regarding its sociolinguistic dimensions. Focusing here on languages where creak is not used for phonemic or allophonic purposes, it has been shown to serve as a prosodic marker and turn-taking signal in e.g. Mandarin Chinese, Finnish, Maori, Japanese, and Spanish [13-17]. Mentions of possible stylistic or socio-indexical uses are few and far between (e.g. in Mandarin Chinese, it may signal sarcasm, [18]; in Tzeltal Maya and Lachixío Zapotec, 'commiserating voice' [19]).

Creaky voice has also been studied in second-language contexts, and there is some evidence for L1-L2 transfer. Kim studied US heritage speakers of Mexican Spanish and L1 US English-L2 Spanish learners and found that the female speakers in both groups used creaky voice as a boundary marker in their Spanish in ways similar to those used in their English (and unlike L1 Mexican Spanish speakers) [20]. Kim ties this result explicitly to the socio-indexical meaning of creak in US English. A complicating factor is that L1 Spanish speakers from across the Spanish speaking world have been found to use creak as a boundary marker [17]. However, in that study it was male speakers who had the highest creak levels, suggesting Kim's results may indeed have a social dimension. Conversely, Gibson et al. found that creaky voice is relatively easily acquired through simple exposure to American English speech without sociolinguistic awareness, as "functionally monolingual" Spanish speakers used it more when repeating English non-words than when repeating Spanish non-words [21]. Finally, Pillot-Loiseau et al. found both transfer and acquisition effects in the interactive setting of conversations between L1 French and L1 English speakers in their respective L2s [22]. Not only did the Anglophone speakers use more creak in their French than the native French speakers did (transfer), the Francophone speakers also used more creak in their English than in their L1 (acquisition).

Creaky phonation has also been shown to have idiosyncratic aspects, to the extent that it is a useful feature for individual speaker identification [16, 23]. There appears to be both inter-speaker variation in the amount of creaky phonation (relative to modal

phonation) and, in bilingual speakers, to what extent a speaker's use of creak is similar across their two languages.

1.2 The present study: L1 Dutch – L2 English

The status of creaky voice in Dutch is largely unclear, particularly outside of speech pathology contexts. It has been described as rare and idiosyncratic [24, 25], but also as something that "often" occurs, though highly idiosyncratically [23]. There are occasional mentions in popular and social media of the term "vocal fry" alongside demonstrations of its use by young female speakers of Dutch, reminiscent of the discourse around it in the US [26].

This paper examines the prevalence of creaky voice in a population of L1 speakers of Dutch who are highly proficient L2 speakers of English. They are students at an international undergraduate college in the Netherlands, where they form a relatively tight-knit community who use English as a lingua franca. Speakers in this community have previously been shown to converge over time on particular phonetic features, both segmental (the realisation of /s/ in their L1 and L2, [27]) and prosodic (L2 speech rhythm, [28]), though not on their use of filled pauses [29, 30]. We study the use of creaky voice in both L1 Dutch and L2 English, to address 1) the dearth of studies on creak in Dutch, 2) questions regarding transfer and/or acquisition of creaky voice in second language contexts, 3) the issue of creaky voice as an idiosyncratic vs. a community feature.

1.3 Expected outcomes

Based on the results from Pillot-Loiseau et al., we expect our speakers to use more creak in English than in Dutch. The ELF setting they are in is largely composed of L2 English speakers, but also contains L1 speakers of American English. Our speakers may also have some awareness of the social function of creaky voice in English. We also expect our speakers to increase their use of creak over time, as this function is acquired. Potentially, we may see some level of convergence (speakers displaying more similar levels of creak over time). Here, the results from previous studies point in different directions: while phonetic convergence between speakers in this corpus has previously been shown, creaky voice may be too idiosyncratic a feature for this to happen. A final potential effect may be an influence from the L2 on the L1, with speakers using more creaky voice in their Dutch over time as a result of English having been their dominant language for a number of years.

2. METHODS

2.1. Corpus

The D-LUCEA corpus [31] contains recordings of 285 students (ages: 17-24) at University College Utrecht, an English-language campus college in the Netherlands. Students carried out a number of speech tasks (reading, monologue, dialogue) in L2 English as well as their L1 (Dutch for the majority of speakers, a variety of languages for the others). The corpus is longitudinal: students from three cohorts were recorded a maximum of five times over their three-year stay. The subcorpus used for the current study consists of the first (year 1, semester 1) and last (year 3, semester 2) recordings of all female speakers from the first cohort whose L1 was Dutch (no other home language growing up). The total number of speakers in this subcorpus is 29.

Each recording contains around four minutes of read speech in English, four minutes of spontaneous speech in English, two minutes of spontaneous speech in Dutch, and the final session also contained read speech in Dutch.

2.2. Automatic detection of creak

Given the difficulties associated with manual annotation of creak [32], there has been great interest in methods for the automatic detection of creak over the years [33, 34]. We use a process based on tracking f0 at 5 ms intervals across the signal with REAPER [35], shown to be the optimal creak detection method for female speakers in an evaluative comparison [36]. Previous use of this method has shown that speakers tend to have bimodal f0 distributions, rather than unimodal ones with creak at the lower tail [16, 37]. The local minimum between the two modes, one for creaky voice and one for modal voice, is the so-called antimode (AM). Using this as a cut-off point allows for each f0 measurement to be classified as either creaky (<AM) or modal voice (>AM), which in turn enables the calculation of creak prevalence per speaker (i.e., the percentage of pitch measurements classified as creaky among the total number of pitch measurements).

2.3. Statistical analysis

We analysed the data using a series of linear mixed-effects models using the *lme4* package in R [38, 39] to predict creak percentage based on language (English, Dutch), style (read, spontaneous), and recording (1 - first, 5 - final). Our data are not entirely balanced, as there is no read speech in Dutch in the first recording. This allows us to compare groups in subsets of the data along three dimensions: within the

spontaneous speech data, we examine the effects of language and recording (model 1); within the English language data, the effects of style and recording (model 2); within the data for the final recording, the effects of language and style (model 3). The final models included speaker as a random effect, and by-speaker random slopes for language (models 1 and 3), style (models 2 and 3) and recording (models 1 and 2). Models 1 and 2 did not include interactions, whereas model 3 did; model comparisons showed that these provided the best fit, respectively.

3. RESULTS

All speakers showed bimodal f0 distributions in both their English and Dutch speech, as shown in Fig. 1. Individual speakers' antimodes were highly similar across their two languages, those for Dutch ranging between 114-164 Hz (M=143, SD=15.1) and those for English between 114-162 Hz (M=143, SD=13.7).

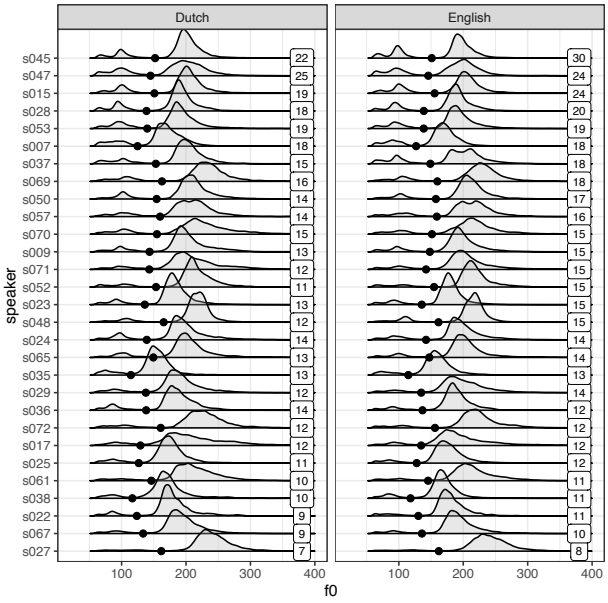

Figure 1: f0 (Hz) distributions, antimodes (●) and creak prevalence (%) by language for all speakers. Speakers sorted by overall creak prevalence.

Mean creak prevalence (over both languages, both tasks, and both recordings) showed a large amount of inter-speaker variability, ranging from 7.67 (SD=1.15) to 26.9 (SD=6.87). In contrast, intra-speaker variation was limited: almost all speakers showed very similar creak prevalence *across* their two languages, with a mean difference of 1.4 percentage points. Inter-speaker variability remained large across recordings, showing no effect of convergence (recording 1: 7-29; recording 5: 9-37).

A linear mixed-effects model predicting creak prevalence in the spontaneous speech data showed an effect of recording, with creak increasing over time (p=.034), but no effect of language. A model predicting creak prevalence in the L2 English data also showed an effect of recording (p=.033), but not of speech style. Finally, a model predicting creak prevalence in the data from the final recording showed an effect of language, with more creak in English than in Dutch (p=.0001), but no effect of style. However, there was a significant interaction between language and style (p=.017): the difference between English and Dutch is found in the read speech style only (with means of 17.5 and 14.6, respectively). The box plot in Fig.2 illustrates the effects of recording, language and style on creak prevalence in the data.

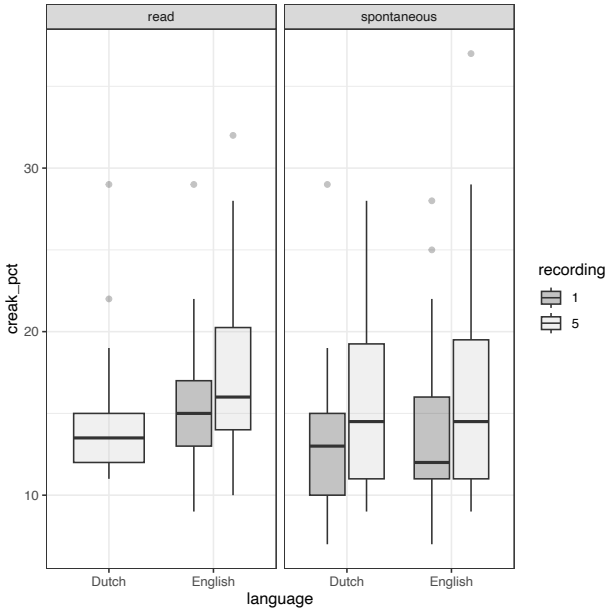

Figure 2: Box plot showing creak prevalence (creak_pct) by language (Dutch, English), speech style (read, spontaneous), and recording (first <1>, final <5>).

4. DISCUSSION

We will start by comparing these results to our expectations. We expected our speakers to use more creak in their L2 English than in their L1 Dutch, because of exposure to creak as a stylistic or socio-indexical feature in English, which – as far as we were aware – is not a function it has in Dutch. We found only very limited effects of language in our data: only in the read speech task during the final recording was there a significant difference between English and Dutch regarding creak prevalence. There are multiple elements to interpreting this result. First, it does not seem that our speakers use particularly low amounts of creak in their English. Although we did not have a clear expectation as to its prevalence, the means and variation displayed by our speakers are comparable to those reported by Dallaston and Docherty [37] for their Australian English speakers, and for those of

American English speakers in some studies [32]. We can, therefore, not conclude that our speakers have failed to acquire English-appropriate creak usage or have limited sociolinguistic awareness. Instead, the lack of a marked difference in creak prevalence between the L2 English and the L1 of our speakers is striking for the opposite reason: the amount of creak used in their Dutch. As noted, very little is known about the creak in Dutch, apart from anecdotal evidence as to its use by young female speakers and claims in older literature that it is rare. The latter notion can easily be put to rest, at least when it comes to the speakers in the D-LUCEA corpus.

The speakers in our corpus are of course not representative of the wider population; they are young, in higher education, and highly proficient in English (a prerequisite for admission to the college). What makes this particular population interesting is that they have previously been shown to display convergence on a number of accent features. The D-LUCEA corpus was collected with the idea of convergence as a "core hypothesis" [31]. This is also partially behind our second expectation, that the use of creak in students' L2 English would increase over time. This appears to be borne out by the data: creak prevalence in English increased between recordings 1 (the students' first semester) and 5 (their last). In fact, creak prevalence also increased in students' spontaneous speech overall, including their Dutch. While it is tempting to relate this to a potential effect of L2 influence on the L1, again the high levels of creak in the participants' Dutch at the outset, and the small difference in creak prevalence between their English and Dutch, casts doubt on this as an explanation. It should also be noted that the effect of recording is relatively small (for both languages, in the order of 3 percentage points).

Finally, actual convergence among speakers, our third expectation, or a potential one, was not observed. The range of creak prevalence among the speakers in the data did not decrease – if anything, it grew slightly wider. This confirms previous studies showing that creak usage is at least partially, and perhaps to a large extent, idiosyncratic, with speakers stable both across languages and over time. While this opens up avenues for forensic research, it puts creaky voice on a par with filled pauses rather than segmental and speech-rhythm features.

This brings us to the three ways this study aimed to contribute to ongoing research. The weight of the evidence on the question of whether creak is mainly an idiosyncratic or community feature seems to firmly land on the former. A second aim of the current study was to engage with the questions of whether creak is a phenomenon easily acquired by L2 learners, or one whose patterns are easily transferred from the L1 to the L2. Given our results, both in terms of the high creak prevalence in Dutch and the stability of speakers across languages and over time, we have to remain agnostic on these questions. The similar levels of creak in individual speakers' L1 and L2 may of course be due to transfer (their use of creaky voice in English simply follows their patterns in Dutch); even more speculatively, these high proficiency, young and female L2 English speakers may have even acquired their use of creak in English and transferred them to their L1. Even if that is the case, however, both its acquisition and its influencing the L1 must have occurred before the recordings in our corpus were made.

Finally, an important contribution of this study is that it is one of very few to look into creak usage in Dutch at all. Our results are to be interpreted with caution given the highly specific nature of our corpus, but they are intriguing even if not representative of wider groups in Dutch society: they suggest that creak may be idiosyncratic, but certainly not rare, among some L1 speakers of Dutch. Creak may simply have flown under the radar of observers, within and outside of linguistics, for a long time. Perhaps its use by Dutch native speakers as a stylistic and/or prosodic feature is not new and not limited to highly proficient English speakers. Going by anecdotal evidence such as mentions on social media platforms, both the term "vocal fry" and its usage (specifically by young female speakers) have increasing visibility. Whether coming from the perspective of forensic linguistics, second language acquisition or sociophonetics, there is ample scope for future study, especially given the tools now at our disposal.

5. REFERENCES

[1] Batliner, A., Burger, S., Johne, B., Kießling, A. 1993. MÜSLI: A classification scheme for laryngealizations. *Proc. ESCA Workshop on Prosody*, University of Lund, 176-179.

[2] Redi, L., Shattuck-Hufnagel, S. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29(4), 407-429.

[3] Keating, P. A., Garellek, M., Kreiman, J. 2015. Acoustic properties of different kinds of creaky voice. *Proc. 18th International Congress of Phonetic Sciences*, University of Glasgow, 2-7.

[4] Davidson, L. 2019. Perceptual coherence of creaky voice qualities. *Proc. 19th International Congress of Phonetic Sciences*, Australasian Speech Science and Technology Association Inc., 147-151.

[5] Davidson, L. 2021. The versatility of creaky phonation: Segmental, prosodic, and sociolinguistic uses in the world's languages. *Wiley Interdisciplinary Reviews: Cognitive Science* 12(3), e1547.

[6] Kreiman, J. 1982. Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics* 10, 163-175.
[7] Laver, J. 1980. *The Phonetic Description of Voice Quality.* Cambridge University Press.
[8] Henton, C., Bladon, A. 1988. Creak as a sociophonetic marker. In: Hyman, L., Li, C. (eds), *Language, speech and mind: Studies in honor of Victoria A. Fromkin*. Routledge, 3-29.
[9] Mendoza-Denton, N. 2011. The semiotic hitchhiker's guide to creaky voice: Circulation and gendered hardcore in a Chicana/o gang persona. *Journal of Linguistic Anthropology* 21(2), 261-280.
[10] Yuasa, I. P. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech* 85(3), 315-337.
[11] Wolk, L., Abdelli-Beruh, N., Slavin, D. 2012. Habitual use of vocal fry in young adult female speakers. *Journal of Voice* 26(3), 111-116.
[12] Davidson, L. 2020. Contributions of modal and creaky voice to the perception of habitual pitch. *Language* 96(1), e22-e37.
[13] Chai, Y. 2019. The source of creak in Mandarin utterances. *Proc. 19th International Congress of Phonetic Sciences*, Australasian Speech Science and Technology Association Inc., 1858-1862.
[14] Ogden, R. 2001. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association* 31, 139-152.
[15] Szakay, A., King, J. 2017. Cross-linguistic transfer effects in bilingual English-Māori voice quality and pitch. *ALS 2017: Conference of the Australian Linguistics Society*, Sydney University, 5-7 December.
[16] Dorreen, K. 2017. Fundamental frequency distributions of bilingual speakers in forensic speaker comparison. MA Thesis, The University of Canterbury.
[17] González, C., Weissglass, C., Bates, D. 2022. Creaky Voice and Prosodic Boundaries in Spanish: An Acoustic Study. *Studies in Hispanic and Lusophone Linguistics* 15(1), 33-65.
[18] Li, S., Gu, W., Liu, L., Tang, P. 2020. The role of voice quality in Mandarin sarcastic speech: An acoustic and electroglottographic study. *Journal of Speech, Language, and Hearing Research* 63, 2578-2588.
[19] Sicoli, M. 2015. Voice registers. In: Tannen, D., Hamilton, H., Schiffrin, D. (eds), *Handbook of discourse analysis*. Wiley, 105-126.
[20] Kim, J. Y. 2017. Voice quality transfer in the production of Spanish heritage speakers and English L2 learners of Spanish. In: Perpiñán, S., Heap, D., Moreno-Villamar, I., Soto-Corominas, A. (eds), *Romance languages and linguistic theory II: Selected papers of the 44th Linguistic Symposium on Romance Languages, London, Ontario*. John Benjamins, 191-207.
[21] Gibson, T. A., Summers, C., Walls, S. 2017. Vocal fry use in adult female speakers exposed to two languages. *Journal of Voice* 31(4), 510.e511-510.e515.
[22] Pillot-Loiseau, C., Horgues, C., Scheuer, S., Kamiyama, T. 2019. The evolution of creaky voice use in read speech by native-French and native-English speakers in tandem: A pilot study. *Anglophonia [online]* 27
[23] van Hugte, T., Heeren, W. 2021. An exploratory study into interspeaker variation in creaky voice in Dutch. *Dag van de Fonetiek*, Utrecht, 17 December.
[24] Collins, B., Mees, I. 2003. *The Phonetics of English and Dutch.* E.J. Brill.
[25] Jenner, B. 1987. Articulation and phonation in non-native English: The example of Dutch-English. *Journal of the International Phonetic Association* 17(2), 125-138.
[26] van Oostendorp, M. 2016. Iedereen doet vocal fry! *Neerlandistiek*. https://neerlandistiek.nl/2016/07/iedereen-doet-vocal-fry/.
[27] Quené, H., Orr, R., van Leeuwen, D. 2017. Phonetic similarity of /s/ in native and second language: Individual differences in learning curves. *JASA Express Letters* 142(6), 519-524.
[28] Quené, H., R., O. 2014. Long-term convergence of speech rhythm in L1 and L2 English. *Proc. Speech Prosody 2014*, ISCA, 342-345.
[29] de Boer, M. M., Heeren, W. F. L. 2020. Cross-linguistic filled pause realization: the acoustics of uh and um in native Dutch and non-native English. *Journal Of The Acoustical Society Of America* 148(6), 3612-3622.
[30] de Boer, M. M., Quené, H., Heeren, W. F. L. 2022. Long-term within-speaker consistency of filled pauses in native and non-native speech. *JASA Express Letters* 2(3), 035201.
[31] Orr, R., Quené, H., van Beek, R., Diefenbach, T., van Leeuwen, D. A., Huijbregts, M. 2011. An international English speech corpus for longitudinal study of accent development. *Proc. Interspeech*, ISCA, 1889–1892.
[32] Dallaston, K., Docherty, G. 2020. The quantitative prevalence of creaky voice (vocal fry) in varieties of English: A systematic review of the literature. *PLoS ONE* 15(3), e0229960.
[33] Martin, P. 2012. Automatic detection of voice creak. *Proc. Speech Prosody 2012*, ISCA, 43-46.
[34] Kane, J., Drugman, T., Gobl, C. 2013. Improved automatic detection of creak. *Computer Speech and Language* 7, 1028-1047.
[35] Talkin, D. 2015. REAPER: Robust Epoch And Pitch EstimatoR. https://github.com/google/REAPER.
[36] White, H., Penney, J., Gibson, A., Szakay, A., Cox, F. 2022. Evaluating automatic creaky voice detection methods. *The Journal of the Acoustical Society of America* 152(3), 1476-1486.
[37] Dallaston, K., Docherty, G. 2019. Estimating the prevalence of creaky voice: A fundamental frequency-based approach. In: Calhoun, S., Escudero, P., Tabain, M., Warren, P. (eds), *Proceedings of the 19th International Congress of Phonetic Sciences*. 532-536.
[38] R Core Team. 2022. *R: A language and environment for statistical computing.* ISBN 3-900051-07-0.
[39] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1-48.