Session 2B Abstracts

Subcategorical representations of corner vowels versus non-corner vowels

Michael C. Stern, Kyle Gorman & Gita Martohardjono
The Graduate Center, City University of New York

Previous research has suggested that the cognitive representations of speech sound categories have gradient internal structure. In particular, category goodness rating tasks have demonstrated that some exemplars are perceived as better or more prototypical than others. Moreover, the best exemplars or prototypes of speech sound categories have been argued to play an important role in speech perception, minimizing the perceptual distance between themselves and neighboring stimuli in the same category like a 'magnet' (e.g. Kuhl, 1991). However, the majority of this research has focused on the corner vowel /i/, and attempts to extend these findings to different categories have had only mixed success (e.g. Sussman & Gekas, 1997; Thyer et al., 2000).

This study extends work on gradient category goodness to one corner vowel (American English /æ/) and one non-corner vowel (Turkish /y/). A matrix of 25 stimuli was synthesized for each vowel space, varying in F1 and F2 by steps of 30 mels. F1 and F2 values for the /æ/ stimulus in the center of the matrix were based on measurements from Peterson & Barney (1952); those for the center /y/ stimulus were based on Radisic (2014). A norming identification experiment confirmed that neither matrix overlapped with another category. Native speakers of American English ($n = 23$) and Turkish ($n = 20$) completed goodness rating tasks where they listened to each stimulus and rated how well it fit in its category on a 6-point Likert scale. The results from participants' native language are reported here.

As seen in Table 1, American English speakers showed a clear pattern of gradient internal category goodness for /æ/. Ratings increased smoothly as F1 and F2 increased, with the most peripheral stimulus (E1) having the highest rating. This pattern was confirmed by significant correlations between goodness and F1 (Spearman's *rho* = .81, $p < .001$) and goodness and F2 ($r = .48$, $p < .05$). Turkish speakers, on the other hand, did *not* show a clear goodness pattern for the /y/ category (Table 2), confirmed by a lack of correlations between goodness and F1 ($r = .19$, $p = .36$) or F2 ($r = .22$, $p = .30$). We also compared the standard deviations of goodness ratings on the /æ/ stimuli versus the /y/ stimuli using a Welch's *t*-test, and found that standard deviations were greater for the /y/ stimuli, suggesting greater variability in responses, although this difference only approached significance [$t(46.01) = -1.99$, $p = .05$].

We argue that the clear pattern in Table 1, compared to the lack of a pattern in Table 2, is due to the fact that the American English /æ/ category is in a corner of the F1-F2 space, while Turkish /y/ is not (it is bordered by the more peripheral /i/; Radisic, 2014). /æ/ goodness, then, can be represented *qualitatively* (i.e., goodness = high F1 and high F2). /y/ goodness, on the other hand, must be represented *quantitatively* (i.e., with specific values of F1 and F2). It is likely that goodness rating tasks measure qualitative subcategorical representations, but are not sensitive enough to measure quantitative representations. This is supported by the fact that the F1 (896 Hz) and F2 (1921 Hz) of the highest-rated /æ/ stimulus (E1) were much higher than those of average productions (F1 = 660 Hz, F2 = 1720 Hz; Peterson & Barney, 1952).

A follow-up goodness rating experiment (in progress) will test American English speakers on a matrix of /æ/ stimuli even more peripheral in the F1-F2 space (E1 in the matrix from the completed experiment will be A5 in the matrix for the follow-up experiment). We predict that goodness ratings will continue to increase as F1 and F2 increase—even as the stimuli stray very far from average real-world productions—supporting the argument that goodness rating tasks measure qualitative but not quantitative representations. Theories that rely on notions of gradient category goodness need to contend with this methodological limitation.

	1	2	3	4	5
A	-0.39	-0.40	-0.64	-0.94	-0.92
B	-0.02	0.09	-0.21	-0.40	-0.60
C	0.22	0.24	0.19	-0.04	-0.40
D	0.37	0.58	0.60	0.25	-0.18
E	0.91	0.87	0.73	0.48	-0.38

Table 1. Mean goodness ratings (z-scored by participant) of each /æ/ stimulus by English speakers. The stimuli increase in F1 downward, and increase in F2 leftward. Darker green indicates higher ratings.

	1	2	3	4	5
A	0.08	-0.05	-0.08	-0.18	-0.24
B	-0.06	0.17	-0.04	0.00	-0.33
C	0.10	0.30	0.05	0.31	-0.20
D	-0.20	0.30	0.07	0.17	0.02
E	-0.18	0.05	-0.09	-0.01	0.06

Table 2. Mean goodness ratings (z-scored by participant) of each /y/ stimulus by Turkish speakers. The stimuli increase in F1 downward, and increase in F2 leftward. Darker green indicates higher ratings.

References

Kuhl, P. K. (1991). Human adults and human infants show a ‘perceptual magnet effect’ for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50*(2), 93-107.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*(2), 175-184.

Radisic, M. (2014). An ultrasound and acoustic study of Turkish rounded/unrounded vowel pairs. PhD thesis, University of Toronto.

Sussman, J. E., & Gekas, B. (1997). Phonetic category structure of [ɪ]: Extent, best exemplars, and organization. *Journal of Speech, Language, and Hearing Research, 40*, 1406-1424.

Thyer, N., Hickson, L., & Dodd., B. (2000). The perceptual magnet effect in Australian English vowels. *Perception & Psychophysics, 62*(1), 1-20.

Prosodic Prominence and High Vowel Lowering in Apurímac Quechua

Sreeparna Sarkar and Irene Vogel

University of Delaware

Like other dialects in the Quechua IIC category (e.g., Cusco, Puno, and Bolivian Quechua), Apurímac Quechua, a southern Peruvian dialect, exhibits a three vowel system with /i, u, /a/ in native words[1]. Mid vowels /e, o/ may also be observed, especially in Spanish loan words, though their status is somewhat controversial. In addition, earlier studies of Cusco[2] and Cochabamba Quechua[3] have shown lowering of the high vowels /i/ and /u/ in the vicinity of uvular stops. This study investigates whether High Vowel Lowering (HVL) applies in Apurímac Quechua as well. In addition to the effect of uvular place of articulation, we also consider the possible effect of prosodic properties on lowering. Specifically, we investigate the possibility that stress and focus, as prominence phenomena, play a role in the manifestation of the high vowels.

Our corpus consists of recordings of one male native speaker of Apurímac Quechua reading real three-syllable CVCVCV words. Given the relative infrequency of /q/ in the onset of Syllable 3 (Syll3), and possible word-final effects, we examined the stops in the first two syllables. In each case, /q/ and /t/ were followed by 10 instances of each of the vowels, /i/ and /u/, in native Quechua words. The stimuli were embedded in two short dialogues, one priming a reading with focus on the target (Focus Condition), the other with focus on a word after the target (Non-Focus Condition); total =160 targets. The structures are illustrated in Table 1; the focused items are bolded. Only the target word appearing in the answer of the dialogues was used for the analysis.

Table 1. Carrier Dialogues

Focus Condition	Q: ¿Imatataq Maria kutiparan kunam illariypi?	A: Maria “**qiquna** ” kutiparan kunam illariypi.
	‘What did Maria say in the morning?’	‘Maria said “**qiquna**” in the morning’.
Non-Focus Condition	Q: ¿Maria “qiquna ” rimapakuranchu tardinpayta?	A: Manam, Maria “qiquna” kutiparan tardinpayta, chaymantaq manaña kaytarimapakuranchu.
	‘Did Maria say “qiquna” **in the afternoon**?	‘No, Maria said “qiquna” **in the morning**, not in the afternoon.

Since prominence often causes enhancement or hyperarticulation of certain acoustic properties, or at least resistance to their erosion[4], we considered HVL in relation to two types of prosody, stress and focus. Specifically, we compared the stressed penultimate Syll2 with Syll1, and the manifestation of the stimuli with and without focus, testing the following hypotheses:

Hypothesis 1 (H1): Apurímac Quechua will exhibit HVL in all conditions.
Hypothesis 2 (H2): Apurímac Quechua will resist HVL in the position of lexical stress.
Hypothesis 3 (H3): Apurímac Quechua will resist HVL under focus.

Confirmation of H1 would mean that Apurímac Quechua (AQ) behaves similarly to the other varieties of Quechua with regard to HVL. Confirmation of H2, H3, or both, would indicate that lexical and / or sentence level prominence preserves the basic high vowels positions.

The recordings were segmented manually using Praat[5]. The vowels were identified by the presence of a consistent periodic waveform and the F1 and F2 values at the mid-point were determined. Figures 1 - 3 show the vowel plots for /i/ and /u/ after the uvular /q/ and alveolar /t/ overall with a generic /a/ for reference, in unstressed vs. stressed syllables (Syll1 vs Syll2), and in focus and non-focus contexts, respectively. As can be seen, there is no evidence of HVL when the conditions are pooled (Fig1). There is also considerable overlap of /i/ in Figures 2 and 3. We do, however, find that /u/ is lower after /q/ than after /t/ in the unstressed Syll1, while there considerable overlap in the stressed Syll2 (Fig.2). Focus prominence does not have a similar effect of resisting HVL: /u/ is somewhat lower after /q/ than /t/ in both focus conditions (Fig.3).

In sum, when our results are pooled, we see no evidence of HVL, at first glance, differently from the findings for other dialects, and disconfirming H1. When we consider stress, however, we see that the unstressed syllable does exhibit HVL, while the stressed syllable resists it, confirming H2. Focus prominence does not lend the same strengthening effect, with both focus conditions showing an effect of HVL, disconfirming H3. It must be noted, however, that the previous studies[2,3] did not specifically control for different prosodic positions, so it is difficult to determine to what extent our findings are (or not) consistent with them. Moreover, the HVL in the other dialects was determined by plots of F1 against F2; but when the vowels were plotted as F1 against F2-F1, HVL was not observed. We may thus conclude that, at least in Apurímac Quechua, HVL is sensitive to lexical stress prominence, with the stressed syllable resisting lowering, but not to focus prominence. With regard to the Quechua varieties, for a more reliable comparison, it would be necessary to make similar prominence distinctions, and use the same vowel plotting method (i.e., F1 vs. F2-F1).

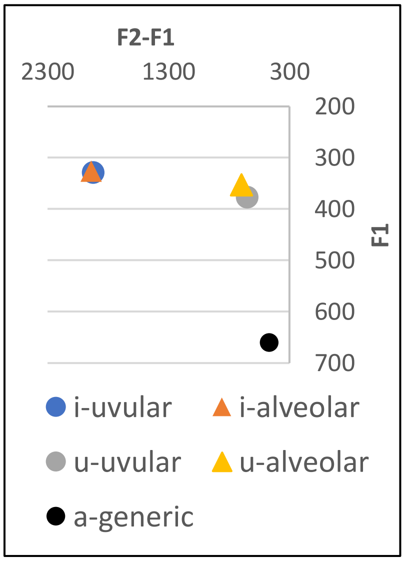

Fig.1. /i/ & /u following Alveolar & Uvular Stops Uvular Stops

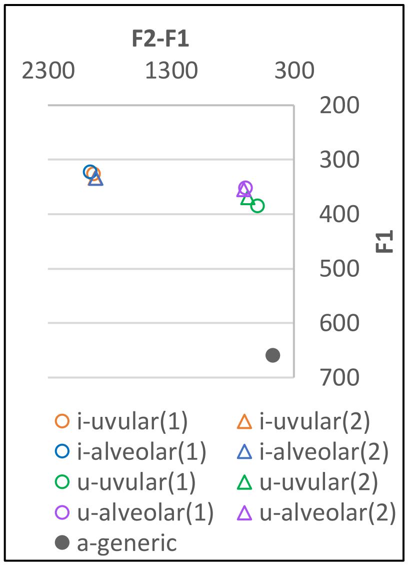

Fig.2. /i/ &/u/ following Uvular & Alveolar Stops / in Syllables 1 & 2

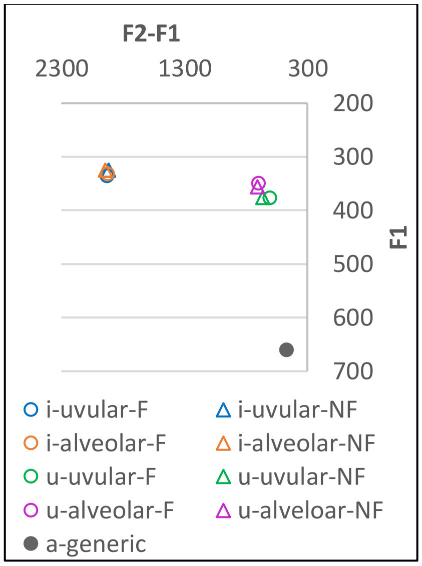

Fig. 3. /i/ & /u/ following Uvular & Alveolar Stops in Focus & Non-Focus Contexts.

References:

[1] Cerrón-Palomino, R. 1994. *Quechua sureño diccionario unificada Quechua-Castellano*, Castellano-Quechua [*Unified dictionary of southern Quechua, Quechua-Spanish Spanish-Quechua*]. Lima: Biblioteca Nacional del Perú.
[2] Molina-Vital, C. 2011. Reconsidering Cuzco Quechua vowels. MS. Rice University.
[3] Gallagher,G. (2015). “Vowel height allophony and dorsal place contrasts in Cochabamba Quechua”, New York University.

[4] Remijsen, B. 2002. Lexically contrastive accent and lexical tone in Ma’ya. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology* 7. 585–614. Berlin: Mouton de Gruyter.

[5] Boersma, P & Weenink, D. 2019. Praat: doing phonetics by computer [Computer pro-gram]. http://www.praat.org/.

Classifying Apurímac Quechua ejectives

Mackenzie Hope Marcinko, University of Delaware, Department of Linguistics & Cognitive Science

I analyze ejective stops in Apurímac Quechua (a variety of Quechua spoken in Southern Peru, henceforth *AQ*) in light of the classification system proposed in Lindau (1984) and Kingston (1985), in which ejectives are considered *strong* or *weak* based on their acoustic properties. I discuss quantitative characteristics of AQ ejectives as determined in an acoustic study and evaluate the suitability of Lindau (1984) and Kingston (1985)'s typology for categorizing ejectives in light of AQ data.

Lindau (1984) and Kingston (1985)'s typology is based on the amount of tension in the vocal folds. *Strong* ejectives (also called *stiff*) result from lengthwise tension of the vocal folds during the glottal closure, resulting in acoustic properties like a long *voice onset time* (*VOT*) and powerful stop burst (Lindau, 1984) (Warner, 1996). Strong ejectives also affect the following vowel, with vowels following strong ejectives exhibiting fast rises to peak amplitude (known as *rise time.* Smaller rise time values are associated with a quicker rise to peak amplitude), tense voice qualities at their onsets, and a falling F0 contour toward the vowel midpoint (Kingston, 1985) (Hargus, 2007) (see Table 1).

Contrasted with strong ejectives are *weak* ejectives (also called *slack* ejectives), which involve little longitudinal tension of the vocal folds. Weak ejectives are characterized by a shorter VOT and less-powerful stop burst than strong ejectives (Lindau, 1984) (Warner, 1996). Vowels after weak ejectives also exhibit slower (longer) rise times than vowels after strong ejectives, as well as more modal voice qualities at their onsets and rising F0 contours toward the midpoint when compared to vowels following strong ejectives (Kingston, 1985) (Wright et al., 2002) (Hargus, 2007) (see Table 1).

Table 1: Ejective typology (Lindau, 1984; Kingston 1985; adapted in Bird, 2002a; Ham, 2007)

	Weak ejectives	**Strong ejectives**
VOT	short	long
Stop burst	quiet	noisy
Rise time	slow to rise	fast to rise
F0 contour	falling	rising
Voice quality	creaky	modal

I examined AQ ejectives for three particular acoustic properties (VOT, rise time, F0 contour). Recordings of ejective and pulmonic stops were collected from sessions with native AQ speaker using Praat and compared across all places of articulation (2 sample t-test assuming unequal variance, alpha at .05) (Boersma & Weenink, 2019). Ejective stops had a significantly longer average VOT than pulmonic stops (42.56 m/s average for ejectives stops; 17.39 m/s for pulmonic stop) The average rise time of vowels following ejective stops (7.74 dB) was also significantly greater than that of vowels following pulmonic stops (4.85 dB). Vowels following ejective stops are therefore slower to rise to their peak amplitude than vowels following pulmonic stops. Average F0 values did not significantly differ between vowels following ejective versus pulmonic stops, so I did not consider F0 in light of the strong or weak typology.

Based on these results, AQ ejectives are neither strong nor weak, having acoustic properties of both strong and weak ejectives. Table 2 shows the three correlates examined in AQ (VOT, rise time, F0 contour) as expressed in other languages. AQ ejectives' long VOT values suggest that

AQ ejectives pattern like strong ejectives, but long rise time values indicate that AQ ejectives are also weak. The binary classification system proposed in Lindau (1984) and Kingston (1985) does not suitably describe AQ ejectives.

Many languages' ejectives do neatly fit within Lindau (1984) and Kingston (1985)'s classification of *strong* or *weak*. Based on the same three properties observed in the current study (VOT, rise time, F0 contour), Hausa has canonically weak ejectives, while Navajo has strong ejectives (Lindau, 1984) **(**McDonough & Ladefoged, 1993) (Table 2). However, our current finding from AQ adds to the growing body of research suggesting that Lindau (1984) and Kingston (1985)'s binary system of classifying ejectives is insufficient. Warner (1996) determined that Ingush ejectives are *mixed*, having acoustic properties in-line with both strong and weak ejectives. Two related Athabaskan languages (Tsilhqut'in and Witsuwit'en) likewise exhibit phonetic characteristics of both strong and weak ejectives (Ham, 2004) (Hargus, 2006). The labels *strong* or *weak* do not adequately describe ejectives in AQ, Ingush, Tsilhqut'in, and Witsuwit'en. I therefore conclude that ejectives ought to be classified along a spectrum, rather than the binary system proposed in Lindau (1984) and Kingston (1985).

Table 2: Ejective classification by acoustic correlates (cells with – denote unavailable or statistically insignificant data)

	VOT	**Rise time**	**F0 contour**	**Overall classification (strong, weak, mixed)**
AQ	ejective > pulmonic	ejective slower than pulmonic	-	mixed
Ingush	ejective > pulmonic	ejective slower than pulmonic	ejectives fall, pulmonic rise	mixed
Tsilhqut'in, Witsuwit'en	ejective > pulmonic	ejective slower than pulmonic	ejective rise, pulmonic fall	mixed
Navajo	ejective > pulmonic	ejective faster than pulmonic	**-**	strong
Hausa	ejective < pulmonic	ejective slower than pulmonic	**-**	weak

Bird, S. (2002a). *Dakelh Ejectives: Evidence for New Ways of Classifying Sounds*. 6th Meeting of the Linguistic Society of America. San Francisco, CA.

Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer. Version 6.0.49, retrieved 26 March 2019 from http://www.praat.org.

Ham, S. (2004). Tsilhquit'in Ejectives: A Descriptive Phonetic Study. Master's Thesis: University of Victoria. Victoria, BC.

Hargus, S. (2007). *Witsuwit'en grammar: Phonetics, phonology, morphology*. Vancouver, Canada: UBC Press.

Lindau, M (1984). Phonetic differences in glottalic consonants. *Journal of Phonetics* (12), 147-155.

Kingston, John (1985). *The Phonetics and Phonology of the Timing of Oral and Glottal Events*. Dissertation, Department of Linguistics, University of California, Berkeley.

Warner, N. (1996). Acoustic characteristics of ejectives in Ingush. *In Proceeding of Fourth International Conference on Spoken Language Processing,* 3, 1525-1528.

Wright, R., Hargus, S., & Davis, K. (2002). On the categorization of ejectives: data from

Witsuwit'en. *Journal of the International Phonetic Association*, *32*(1), 43-77.

NO NEED TO YELL: A prosodic analysis of writing in all caps

Maria Heath • University of Minnesota

In this research, I present a novel, production-based methodology for probing the relationship between the non-standard orthography of social media and prosody. Specifically, I demonstrate that the prosodic realization of the all caps pattern systematically includes higher volume as well as higher pitch and increased tempo. Although many laypeople and Computer Mediated Communication scholars (McSweeney 2018, McCulloch 2019) share the intuition that writing in all caps amounts to “yelling”, my data highlights the nuanced prosodic capabilities of non-standard orthography.

Little analytic work has been done to date on the all caps orthographic pattern as used on social media and in other synchronous written communication. Analysis of CMC data that includes non-standard orthography typically relies on authors’ intuitions to interpret the meaning of the pattern. All caps is often associated with anger (or "flaming"), but it is clear from the examples below (from a corpus of tweets I collected in 2016) that its actual meaning extends beyond this use, and it has largely remained elusive. Some have proposed that it serves to increase emotive strength (Zappavigna 2012), a theory which can explain examples (2) and (3) as well as (1).

(1) Anger: THERE ARE ALREADY LAWS ON THE BOOKS CONCERNING GUNS BUT THEY ARE NOT BEING ENFORCED! MORONS!

(2) Excitement: IM OFFICIALLY IN CALIFORNIA AND IT IS BEAUTIFUL

(3) Intense desire: I NEED TO DO THIS

While this explanation fits with use of all caps in contexts of anger or excitement, it cannot account for the absence of the all caps pattern in tweets indicating intense sadness or boredom and its rejection by readers in such contexts (Heath 2018).

Instead of trying to forge a direct link from orthographic form to meaning for all caps text, I instead propose that it is a way to encode specific prosodic features associated with higher states of emotive arousal. This approach neatly explains why sadness and boredom, even in their strongest forms, cannot be expressed by all caps, since they are low-arousal emotions which cannot be heightened with increased F0 or intensity. On the other hand, anger and excitement, for example, are high arousal emotions and can be heightened in this way.

To investigate whether social media users systematically associate all caps text with particular prosodic patterns, I recorded 20 people reading aloud a series of 48 stimuli tweets, some written in all caps and others with standard capitalization, as exemplified in Figure 1.

Figure 1: Example all caps stimulus and standardized stimulus

Quinicine Houndy
@quinicinehoundy180
WHY IS NO ONE TALKING ABOUT THIS?
5:52 PM - 28 June 2017

Quinicine Houndy
@quinicinehoundy180
Why is no one talking about this?
5:52 PM - 28 June 2017

I asked participants to read the tweets aloud as they best imagined the author of the tweet intended, but did not mention intonation specifically. I used a linear mixed effects model to compare means and ranges of F0, intensity, and tempo across the two conditions of resulting recordings.

For both F0 and intensity, the means for each participant were higher when they were reading stimuli written in all caps compared to stimuli written in standard orthography. F0s for the standard condition were lower than the all caps condition by 11.68 Hz ± 2.43 ($p<0.5$) and intensities for the standard condition were lower by 3.06 dB ± 0.55 ($p<0.05$). In other words, participants tended to speak all caps text louder and with higher pitch. The intensity range was also higher for stimuli in all caps ($p<0.05$), with the standard condition lower by 1.69 dB ± 0.48, indicating that the increased mean intensity may not have been simply a shifting up of intensity across the whole utterance, but an increase in intensity peaks. However, no significant difference was found for F0 range ($p=0.800$), suggesting that the increase in mean F0 was more likely an overall upward shift of F0 rather than an increase of peak F0 height. The tempo data has not yet been fully analyzed, but initial observations suggest that these results may be more complex, with tempo increasing or decreasing depending more on the specific content of the stimuli.

My results suggest that prosody is a critical element to understanding how the all caps pattern gets interpreted on social media. I propose that a deviation from an expected orthographic form is noticeable and must indicate some additional piece of information. Prosodic interpretations are a good candidate for this additional information, considering that silent reading typically involves a mental “realization” of a text as if one were reading it out loud, a phenomenon known as “silent orality” (Soffer 2010). While standard English orthography doesn’t provide many explicit clues about intended prosodic realization, this age of synchronous written communication calls for a more robust system for indicating emotion, attitude, and even semantic structure through creative prosodic orthography. The all caps pattern is only one of many orthographic patterns which do prosodic work, and my research sets the scene for further analysis of such patterns and an increased understanding of the role of prosody in silent reading.

References

Heath, Maria. 2018. Orthography in Social Media: Pragmatic and Prosodic Interpretations of Caps Lock, In Proceedings of the Linguistic Society of America Annual Meeting, Jan. 2018, Salt Lake City, UT.

McCulloch, Gretchen. 2019. *Because Internet: Understanding the New Rules of Language*. New York: Riverhead Books.

McSweeney, Michelle. 2018. *The Pragmatics of Text Messaging: Making Meaning in Messages.* New York: Routledge.

Soffer, Oren. 2010. ‘‘Silent Orality’’: Toward a Conceptualization of the Digital Oral Features in CMC and SMS Texts. Communication Theory 20. 387–404.

Zappavigna, Michele. 2012. *Discourse of Twitter and Social Media*. Bloomsbury: Bloomsbury Publishing.