Culture and Interdiscursivity in Korean Fricative Voice Gestures

This paper explores the cultural significance of a type of audible gesture in Korean speech that I call the Fricative Voice Gesture (FVG). I distinguish between two forms of this gesture: the reactive FVG, which serves as a self-standing utterance that signals personally felt intensity, and the prosodic FVG, which can be superimposed upon an utterance as a form of intensification. Based on an ethnographically informed analysis of interviews, Christian sermons, and advertisements for soju, a Korean spirit, in South Korea, I view the interdiscursive link between reactive and prosodic FVGs in terms of the ongoing cultural revalorization of the sound shape. I focus in particular on the shift from harsher to softer FVGs—and their omission altogether—according to different, but related, paradigms of social differentiation such as class, gender, and age. [voice, gesture, prosody, intensification, korean, South Korea]

There is always something about the voice that must be ascribed to the social background, precisely as in the case of gesture. Gestures are not the simple, individual things they seem to be. They are largely particular to this or that society. In the same way, in spite of the personal and relatively fixed character of the voice, we make involuntary adjustments in the larynx that bring about significant modifications in the voice.


This paper is about a type of audible gesture in Korean speech. I call this sound the Fricative Voice Gesture (FVG), and I begin with two examples of how this sound can operate. In the first example, the FVG serves as a self-standing utterance; in the second example, the FVG serves as a prosodic element that is superimposed upon an utterance. I spend the remaining portion of the paper discussing how these two types of FVGs are related as culturally significant aspects of Korean communication.

While doing fieldwork in Seoul during 2008 and 2009, I sang with a Presbyterian church choir. After a rehearsal or a service, the choir members often would gather together for a meal paid for by one of the deacons or some other senior member of the choir. In late August, after enjoying just such a meal of soybean-paste stew (toenjang tchigae) near the church, I walked out of the restaurant with some of the other men, stood in the middle of the small street, and began sucking small bits of rice out of my teeth, just as the others were doing. Any Korean will recognize such a scene of well-fed men standing in the middle of the street, often with toothpicks, doing this sort of thing. When Yongch’an, a fellow tenor in his mid 30s, noticed me doing this, he took my arm and said, “When you go back to the U.S., your wife will...
“hate you.” “Why?” I asked. “Because you are becoming a Korean ajossi!” he said. Ajossi is a common term for a man of marriageable age—a grown-up guy—but it also has become a kind of pejorative term used to describe a type of Korean man who is loud, abrasive, not refined, and not modern. So I responded, “And what if I make this sound?” I then emitted a harsh consonant-like sibilant sound that I had heard many Korean men make after taking a drink of soju, a widely-consumed Korean spirit: “[kʰiː sứːiː].” Yongch’an laughed and said, “She will send you back to Korea!” Then he waved his hands in front of his face and voiced her virtual response: “Nomu sikkirōwō!” (Too noisy!). Yongch’an was sure that this harsh sound, which I call a reactive FVG, would be too much for an American woman to bear. “Anyway,” he followed, “now we say [kʰjaːː],” producing a softer, brighter-sounding version of the sound I had just made.

Later on that year, on a Saturday afternoon in October, I enjoyed a seafood buffet with the choir conductor and two other members of the choir. The conductor had invited us for lunch to celebrate the success of the choir’s fall concert, which had taken place a few days earlier. The conductor had also brought her youngest son, Sehyŏn, to dine with us. Sehyŏn is a ham. He is friendly and loquacious with new people, enjoys singing and performing for strangers, and has a particularly keen ability to mimic the voices of others. He did a biting rendition of an ajossi speaking Korean with an American English accent—which I did not take personally, of course. Figure 1 features two pictures of Sehyŏn making comical use of a king-crab claw from the buffet.

Needless to say, Sehyŏn enjoyed being a focus of attention and was already, at four years old, quite skilled at maintaining an audience. During lunch, Sehyŏn became fascinated by my recording equipment and camera. He repeatedly asked me to record his various performances, which included stories, mimicry, and songs.

One of his stories (Excerpt 1) was about an angel who saves some soldier ants from being eaten by a king ant. Sehyŏn displayed his developing virtuosity by accentuating the textual poetics of his story through visible gesture, as can be seen from the video stills included to the right of the transcript. Sehyŏn’s wide grins (1, 8) served as narrative brackets for opening and closing the story. As he provided the setting for the drama to come, he leaned into the camera to secure my attention (2). When introducing the “very big” king ant, his face expressed surprise and fear while his hands trembled (3). Then as he described the ants being eaten, he mimed the process of scooping the small ants into his mouth (4). After introducing the angel with another smile (5), he directed his eyes away from the camera and shook his finger as if he himself were driving away the monster (6). Finally, when describing how the
angel took care of the ants, he pressed his cheek to the back of a chair and caressed it (7).

Alongside these gestures and other prosodic manipulations, Sehyŏn uttered a quick, harsh sound resembling the FVG that I had produced for Yongch'ăn (see above). But instead of uttering it on its own, he did so while pronouncing the first [a] in the word kapchagi (‘suddenly’) in line three (noted as [a’]). Rather than emitting the FVG as a self-standing reactive utterance as I had done, Sehyŏn superimposed it as a prosodic element over his utterance as a kind of “rhetorical accent” (Newman 1946 qtd. in Bolinger 1961) to signal the intensity of the king ant’s actions. The fact that he made the sound when he introduced the king ant, but not when he used the same adverb, kapchagi, to introduce the angel, also suggests that it played a role in marking an important poetic contrast in the story. The significance of this contrast will become clearer as I proceed with my analysis.

My approach in this paper is to focus on the “sound shape” (Jakobson and Waugh 1987[1979]) of FVGs across these two types of use. I argue that the prosodic FVG’s function as an intensifier in speech is linked to the reactive FVG as a legitimate form of expression locatable in prototypically male-centered drinking events. After taking a first sip of soju, drinkers emit an isolated, non-prosodic FVG as a form of expressive reaction to the experience of consuming the alcohol. Korean drinking events have long served as a ritual site of the age- and sex-based stratification of Korean society, and the reactive FVG is a recognizable index of the privileging of males—in particular, the ajossi—as the center of this event. On the other hand, when an FVG is superimposed onto the phonological material of speech as a form of prosodic intensification, it also serves as a speaker-centering index of the authority to express personally experienced intensity. That is, it signals metasemantically the intensification of some stretch of denotational content for a speaker and, crucially, signals metapragmatically (Silverstein 1976, 1993) the contextual legitimacy of this intensity’s expression. My basic point is that we can understand the prosodic FVG in terms of the reactive FVG and vice versa.

I demonstrate this, first, by accounting for the way in which my informants linked the use of prosodic FVGs to categories of class and age. Then, I look at how the perceived appropriateness of FVGs in speech was characterized as a function of the relative authority of a speaker vis-à-vis other speakers. I explore these ideological framings by comparing the sermonic styles of two pastors, one Pentecostal and one Presbyterian, in order to show how the use or suppression of FVGs, and the communication of personally experienced intensity more generally, figure into broader institutional registers of communication.

Finally, I return to the reactive FVG in soju consumption to show how the stratification and ideological framing of prosodic FVG usage parallels a cultural revalorization of reactive FVGs. As the soju industry increasingly markets its products to women, the harsh reactive “male” FVG has given way to a softer, more gender-neutral version (as in Yongch’ăn’s revision of my FVG described above). As this is taking place, the widespread use of prosodic FVGs—available to all speakers, according to context—is also increasingly construed by younger, wealthier generations as a behavior of old or lower-class people. Thus, there is a softening of both types of FVGs in accordance with different, but related, paradigms of differentiation.

I deal with prosodic form in terms of a broader semiotics of communication and culture, paying special attention to sonic iconicities and their communicative functions across seemingly disparate types of semiotic events. In this way, I can suggest that the interdiscursive relationship between reactive and prosodic FVGs involves a functional rank-shifting between the plane of utterance types and the plane of prosodic types, i.e., the same type of sound shape shifts from serving as a reportable utterance to serving as a paralinguistic framing of an utterance. I conclude by returning to Sehyŏn’s story about ants and angels to show how his poetic use and omission of the prosodic FVG diagrams the shifting place of FVGs in broader South Korean society.
Excerpt 1: Of Ants and Angels (photo by the author)

1. 어느 날, 응, 어, 병정들이 가고 있는데
   Önū nal, ūng, ō, pyōngjōng-dul-i kago innunde
   One day, um, uh, some soldiers were moving,

2. 개미들이 땅 속을 가고 있는데
   kaemi-tūl-i ttang sog-ūl kago innunde.
   some ants were moving inside the earth.

3. 갑자기 엄청 큰 왕—왕개미가
   K[a]pchagi omch'ōng k'un wang—wanggaemi-ka
   Suddenly a very big king—king ant

4. 조그만 개미[를] 다 잡아먹은 거예요.
   chogūman kaemī-[rūl] ta chabamōgūn kōeyo.
   grabbed the little ants and ate them.

5. 갑자기 천사가 나타나가지고
   K[a]pchagi ch'onsa-ka nat'anagajigo
   Suddenly an angel appeared,

6. 괴물을 몰리쳐가지고
   koemul-ūl mullich'yōgajigo
   and drove back the monster, and

7. 어, 천사가, 어, 개미들을 착하게 키워 왔어요.
   ō, chōnsa-ka, ō, kaemi-tūl-ūl ch'akhage
   kīwō chuōssōyo.
   uh, the angel, uh, kindly raised the ants.

8. 이게 끝이이요.
   Ige kkūch'ieyo.
   That's the end.
I. Fricative Voice Gestures

Fricative voice gestures are emissions of sound produced when air passes through sites of frication along the supralaryngeal vocal tract. In my observation, the sound can range from breathiness or a strong whisper to the rougher, voiceless sounds that one might hear when someone clears his or her throat. In its prosodic usage, the FVG occurs on the initial vowel of adverbs and, less frequently, adjectives, and is often coupled with lengthening. When a prosodic FVG is emitted, the vowel is no longer phonated from the vocal cords. Rather, as air passes through the vocal tract, the locus of frication itself becomes the primary site of sound production, while the articulators farther up the vocal tract remain intact, so that the vowel is still perceived as a vowel, even when lengthened. Thus, in prosodic FVGs, as in fricative vowels, consonant-like phonetic features serve phonemically as vowels (see Ladefoged and Maddieson 1996).

Take, for example, the Korean adverb chŏngmal, glossable as ‘really’ or ‘truly.’ In Figure 2 below, I give the transcription for a phrase uttered with “neutral” intonation. Below it, I provide an example of the same word with lengthening and rising intonation, a commonly heard form of intensification. Finally, I provide an example with intensificatory lengthening and an FVG, which informants described as stronger than rising intonation with lengthening.

If isolated from the spoken word and examined phonetically for its features, the FVG in Figure 2, [ʌx], probably would appear to be the consonant [x] (or [xʷ] if labialized, as in the rounded form of the same Korean vowel, [ɔʷ]). However, when used prosodically, the sound is articulated and interpreted as a segmentable vowel in accordance with the distinctive features of Korean. In my data, the consonantal sound, produced by partial closure of the vocal tract, is superimposed prosodically over, and replaces, vocal-cord phonation, an inherent feature of the particular phoneme (see Jakobson, Fant, and Halle 1969:13; see also Fónagy 1976). So if “prosodic features . . . can be defined only with reference to a time series” (Jakobson, Fant, and Halle 1969:13), the FVG is contrasted not with the phonological material immediately preceding or following it, but, rather, with vowels in poetically sequential, not phonologically sequential, places (i.e., the vowels of the first syllables of other adverbs, as in Sehyŏn’s contrasting tokens of “kapchagi”; see Jakobson 1960).

Following Sapir, I treat this prosodic form as a kind of gesture. In his 1927 paper “Speech as Personality Trait,” Sapir made the following claim: from a “general point of view, voice may be considered a form of gesture. If we are swayed by a certain thought or emotion, we may express ourselves with our hands or by some other type of gesturing, and the voice takes part in the total play of gesture” (1927:895). Later, in his 1931 entry “Communication” for the Encyclopaedia of the Social Sciences, he elaborated further: “Gesture includes much more than the manipulation of the hands and other visible and movable parts of the organism. Intonations of the voice may register attitudes and feelings quite as significantly as the clenched fist, the wave of the hand, the shrugging of the shoulders, or the lifting of the eyebrows” (1931:78). Gestures, in this sense, can be seen as multimodal paralinguistic signs that are related indexically to, synchronized with (Bolinger 1986:198), and superimposed upon some stretch...
of uttered denotational text in discursive interaction. The timed simultaneity of gesture gives it its indexically calibrated metacommunicative functionality.

The notion of the voice as a medium for gesture has been used for diverse purposes.11 George Herbert Mead’s (1934:46) notion of the “Vocal Gesture” dealt with the voice’s potential to emit “significant symbols,” acts of social behavior that would serve as stimuli to others, would be imitated and responded to, and would be meaningful to emitter and receiver alike. Norman McQuown (1954:56) used the phrase “vocal gesture” to describe “any phonetic material produced by the vocal organs which does not fit comfortably into established phonemic structures.” Phonetician Maurice Grammont (1946[1933]:413–4) used the term “gestes articulatoire” (‘articulatory gestures’) to denote articulo-semantic iconisms in which the mouth shapes resemble the things they are supposed to represent. Roman Jakobson used the terms “gestes vocaux” (‘vocal gestures’) (1990[1939]:308) and “sound gestures” (1990[1960]:354) to describe a “natural” onomatopoeia that makes up the prephonological (or pre-phonologized) babble of infants. Jakobson viewed these “vocal gestures” as existing prior to their socio-linguistic “taming,” before “the phonetic richness of the babbling period thus gives way to a phonological limitation” (1990[1939]:296).12 Dwight Bolinger (1986:198) also believed that paralinguistic gestures of the vocalic (audible) or kinesic (visible) sort likely emerged diachronically from a state of absolute motivation (e.g., a histrionic cry or a defensive posture against attack) and eventually were conventionalized into “the gestural complex of which intonation is a part.”13

Rather than starting from a presumption of the diachronic evolution from some state of absolute motivation to one of conventional, arbitrary prosodic form, this paper examines the relative motivation of FVGs as a particular prosodic feature in Korean speech in relation to a broader contemporary semiotics of audible expressivity. I use the term “gesture” to refer to the semiotic function of paralinguistic signals that relate an uttered text to its context by “commenting upon” both the denotation and a framework of its appropriateness and interpretability (see Gumperz 1982 on “contextualization cues”).

The indexical effectiveness of gestural frication in Korea is bound to the extent to which it is culturally understood to be iconic of its target, i.e., representational of intensity. But FVGs are neither accounted for in Korean textbooks nor represented in han’gul, Korean’s morpho-phonemic writing system. So while Korean language learners often perceive FVGs as one of the most distinctive aspects of the Korean speech that they hear, they usually are not aware of what the gesture means, nor even aware that frication is operating as a kind of gesture in speech. A Japanese acquaintance of mine in Seoul once asked me to explain the “harsh” and “rough” sounds that she heard “older” Korean speakers making. She assumed that the sound was simply part of a generational sociolect with no particular communicative function. A couple of years earlier, when I was studying Korean in Seoul during the summer of 2006, I myself had asked a classmate of mine, a second-generation Korean American from Washington, DC, about FVGs. This classmate did not immediately know what I meant, so I tried to produce the sound. She responded, “Hey, watch it! That sounds just like my mother!”—as if I had been making fun of the sound. When I asked what the sound was, she responded, “I don’t know. It’s just something she and her friends do. It’s so embarrassing.”

Although non-native speakers and heritage learners of Korean often comment on this sound, native Korean speakers, for the most part, are not immediately aware of the sound’s distinctiveness to outsiders. In my interviews with Korean informants, there was a fairly clear division among groups as to how recognizable the sound was as a segment (or dimension) of prosodic speech. Often, its gestural function did not seem noteworthy upon reflection. Older speakers (e.g., senior church members) and those from a lower socioeconomic background (e.g., taxi drivers and hairdressers) often had more difficulty “hearing” the sound as such, even when I performed the three different versions listed in Figure 2. Many heard only increased passion, excite-
ment, or emphasis, i.e., they recognized the importance placed on a speaker’s utterance without being able to segment or reproduce the source of intensification itself.14

I found that younger speakers, as well as those from a higher socioeconomic background, far more quickly identified the prosodic sound as isolable and, in so doing, often distanced themselves from its use—even if I had heard them make such sounds themselves. In fact, many of them readily identified stereotypical persons, exemplary others (Hastings and Manning 2004), who used FVGs (e.g., people working at a market) as well as known individuals. For example, I attended an expensive dinner of “fusion” Korean cuisine with the members of the church choir mentioned above (a deacon footed the bill). The church has approximately 70,000 members and is located in the upper-class Apkujong neighborhood of Seoul. I knew these people well by this point and was in a position to know something about their views about one another’s social and economic status. When I explained that I was interested in FVGs as a kind of prosodic intensification (I did not call them “FVGs” at the time), the young woman to my right, whom I knew to be the daughter of a wealthy business man, quickly looked toward another young woman at the table, whose mother worked in a flower shop and was the sole earner for the family, and said, “Yes, just like her!” At first, the target of this observation (who had, incidentally, become a good friend of mine) did not understand what was being said about her. However, after further conversation with me, she came to hear the FVGs in others’ speech and to recognize them in her own speech; she even reported trying self-consciously to control them in particular situations. At the same time, she helped me better to implement them in my own speech, and we began to use FVGs as a personal inside joke (see Labov 1966, 1972, Trudgill 1972 on in-group or “covert” prestige).

Even without awareness of segmentability, however, speakers still demonstrated a sense of when FVGs were and were not appropriate. I elicited these responses by asking whether one kind of person could talk “this way” to another kind of person, and then produced an utterance using an FVG. The perceived appropriateness of FVGs generally came down to instances when speakers were, or had some claim to be, authoritative centers of attention during interaction. This applied to the elderly, mothers telling bedtime stories, friends among friends, and grown men in general. For speakers deemed ineligible actively to command the attention of their interlocutors because of their title or status (e.g., a student speaking to a professor or a young person asking directions from an older person), FVGs were construed as rude, abrasive, disrespectful, rough, and generally uncommon—unless there was some special intimacy between the speakers.

Judgments about the appropriate use of FVGs in a person’s speech are also judgments about the appropriateness of the person’s social action. Thus, FVGs are not merely part of a first-order local framework of politeness (cf. Brown and Gilman 1960). Like speech levels and honorifics, FVGs figure into a higher-order system of generalizable social relations and attributes (Silverstein 2003). Fricative Voice Gestures in particular instantiate a cultural system of uneven access to communicative opportunities for—indeed, registers of—emphatic and emotive expression (cf. Irvine 1990).15 In the following section, I discuss one example of how the communication of personally felt emotional intensity is licensed.

II. Intensification and Authority

A great example of a virtuosic user of FVGs is David Yonggi Cho, founder and former head pastor of the largest Protestant congregation in the world. Headquartered in Seoul, the Yoido Full Gospel Church (Yoido Sunbogum Kyohoe) has a self-reported membership of 830,000, with a main chapel that seats 12,000 and satellite churches and chapels spread around Korea and the world. The church is known for ministering to the working classes of Korea and beyond, having achieved its growth in part by giving women in the church unordained positions as leaders of “cell groups.”
In a sermon titled “David’s Secret of Defeating Goliath” (May 25, 1997), Cho encouraged the members of the congregation to proclaim their faith “boldly” in order to achieve miracles on Earth. He used St. Peter as an example of initial boldness in faith (Peter eventually falters). In Excerpt 2, the lengthened FVG occurs on the initial [a] of the complex adverb *tamdaeham-úro* (‘boldly, enthusiastically’).

**Excerpt 2: Intensity of Speech, Boldness of Faith**

Cho: 그리고 베드로가 나중에 담대함으로

*Kúróna paèdúro-ka najung-ae t[aːːːː]mdaeham-úro*

*But afterwards, when Peter* **boldly**

*pressoju* 잠은 손을 막 펴어 바렸을 때

*paetchón-e chabún son-úl ttak tteó póryóssúl ttae*

*pulled his clasped hands suddenly from the side of the boat,*

그는 물위에 서기 시작한 것입니다!

*kūn-nun mul-wi-e sógi sjakhun kósimnida!*

*he started to stand on the surface of the water!*

Aud.: 아멘!

*Amen!*

Cho: 여러분—

*Yóribun—*

*Everyone—*

The spectrogram in Figure 3 shows the sustained sibilant sound that occurs in place of normal voicing on the [a] vowel of *tamdaeham-úro*. If isolated from the spoken phrase, this graphic representation would appear phonetically as a consonant.
As Cho told the story of Peter, he walked out to the right side of his pulpit and grasped it as if he himself were Peter, holding onto a boat. Then, on the adverb “suddenly” (ttak), he released his hands and raised them in the air. Then, he completed the phrase and returned to his original place behind the pulpit. Cho thus performed kinesically his own denotational third-person description of Peter, thereby inhabiting Peter’s story while also narrating it. Cho’s own personal “Full Gospel” theology posits that boldness of faith results in charisms from God (Harkness 2010a), i.e., Cho too had “felt” the boldness of faith that Peter had felt. Cho’s prosodic pacing and intensification—particularly the FVG on “boldly” (tamdaeham-iro) and the overall increased stridency of his voice—cued the congregation to respond to his cadence with an “amen.”

From behind the pulpit, Cho followed by addressing the audience directly as “everyone” (yordobun), shifting from a third-person description of Peter to a second-person address to his audience. The third-person account drew on biblical authority as well as Cho’s own intense experience of miracles through personal boldness of faith—laminating a biblical account onto a biographical account—which authorized the sermon’s conical broadcast model of communication and Cho’s position at its apex (Figure 4).

Cho’s FVG did not merely intensify the denotational content, but also intensified the performative event itself, “calling forth special attention to and heightened awareness of both the act of expression and the performer” (Bauman 1977:11). In this way, the legitimately delivered FVG draws a listener’s focus to the denotational text being spoken by signaling not just the intensity of the narrated event, but also of the speaker’s investment in this intensity—and, by extension, the value of the speaker’s sentiments as warrant to command another’s attention.

The example of Cho’s sermon is a straightforward instance of what Austin (1962) called “illocutionary force.” Linguistic anthropologists have argued that the notion of “illocutionary force” describes contextually based entailing indexicals that invoke higher-order ideological systems for their power to induce further social effects from discursive interaction (e.g., Silverstein 1979:215, 222ff.; 1998:142n10). The fact that much of prosody’s effectiveness in use, phonological intensification notwithstanding, is contextually based has long been established (Ervin-Tripp 1976; Gumperz 1982; Bolinger 1986, 1989). The ideological dimension of prosody and its gestural components have not received the same attention, however. Hence one of our projects here is to understand the FVG’s illocutionary force in the “performance” of a speaker’s communication of personally experienced intensity. This question becomes more salient when it is contrasted with a denominationally differentiated speech register that emphasizes solemnity and restraint.
III. From Intensity to Denotational Focus

The example of the Pentecostal preacher David Cho shows how the prosodic FVG and its expressive intensity can be used in the self-anchoring performance of a charismatic religious leader. But not all Korean preachers are the same. Somang Presbyterian Church, where I sang as a choir member during my fieldwork, is a particularly stark example of the institutionalization of vocal normativities of middle- and upper-class Koreans in Seoul (Harkness 2010b). Known as the “intellectual church” in Seoul, Somang’s services are solemn and quiet, and head pastor Kim Chich’ol conducts almost the entire sermon in a hushed voice. He and the church stand as a kind of stylistic opposite of David Cho and his “full-throated” Full Gospel Church. If pastors at Somang Church use frication to mark important elements in their sermons (as illocutionary technique), it usually is produced as a strong whisper, not as the harsher sounds from constriction in the vocal tract that I have been discussing. Furthermore, it generally functions as a form of conceptual emphasis or highlighting (rather than coded emotionality or intensity), which, importantly, can be applied to any portion of an utterance—not just the first vowel of adverbs and adjectives.

Below is an excerpt from a sermon given by a junior pastor at Somang Church, Roh Hanggyu, recorded in Seoul on May 18, 2008. The sermon was based on Mark 5:21 (International Bible Society 1973): “When Jesus had again crossed over by boat to the other side of the lake, a large crowd gathered around him while he was by the lake.” Roh explained that the point of the verse was that Jesus had returned—that he had crossed to the other side again (tasi). When Roh first pointed this out, he repeated the adverb tasi in a strong whisper. Later in the sermon, Roh returned again to the word tasi, this time explaining its importance in the passage for the spiritual lives of Christians: Jesus not only did come again, but also would come again. As he explained this, the whisper reserved for the word tasi “leaked” into the larger phrase of which it was a part, constituting the overall voice quality of entire stretches of the utterance. This larger phrase was Pastor Roh’s conceptualization of the word’s importance in the Bible verse. This is illustrated in the passage below, in which the whispered segments are bolded and underlined.

Excerpt 3: Whisper and Denotational Emphasis

1. 근데 거기에는 오라해서도 아니오.
   Künde kōgi-e-nun orahaesō-to aniyo.
   But it was not because they were told to come there.

2. 먼저 가서 기다린것도 아니었고, 누가 다른 사람들이.
   Mōnjō kasō kidarin-kōt-to aniottgo, nuga tarūn saram-tūl-i
   It also was not the case that they went there first and waited, those other people.

3. 예수님께서 먼저 가실 계획을 가지시고 다시 가셨다. 다시.
   Yesunim-kkesō mōnjō kasil kyehoek-ūl kajisigo tasi kasyotta. Tasi.
   Jesus first had the plan to go [to them] and went again. Again.

4. 어, 예수님께서 우리에게 다시 찾아오신다.
   O, yesunim-kkesō uri-ege tasi ch’ajaosinda.
   Uh, Jesus seeks us again.

5. 다시 찾아오신다.
   Tasi ch’ajaosinda.
   Seeks again.

6. 예수님께서 다시 찾아오신다.
   Yesunim-kkesō tasi ch’ajaosinda
   Jesus seeks again.
Nowhere in the sermon did Roh produce a prosodic FVG in the manner of David Cho. Instead, a whisper recurred throughout the sermon. Note how the whisper is not applied to the verb phrase from the Bible verse, but rather to Pastor Roh’s own paraphrase and elaboration of the importance of the biblical text. As the whisper moved from just the adverb to the whole verb phrase, it “leaked” out as a form of illocution based on denotational focus. While FVGs often can take the form of a whisper, Roh put the sound shape to use in a way that did not cue up a scale of personally felt intensification; rather, it highlighted the hermeneutical significance of the concept for understanding the Bible verse as a whole. I repeatedly observed this technique of explicit explanation and conceptualization in the sermons of Somang Church’s pastors, who often drew on literature and philosophy to clarify aspects of biblical verse, making the sermon approximate a university lecture (Kim Chich’ol was a college professor prior to taking the position as Somang’s head pastor).

Furthermore, the whisper as a paralinguistic marker of conceptual emphasis is related to the broader verbal style of church services. Protestant churches in Korea are known for their emphasis on tongsonggido, individually vocalized prayers uttered in a group setting. Roh’s whispers, superimposed over segments of conceptual importance, are related to the sober, whispered tongsonggido of Somang’s congregants at Wednesday-evening prayer service, just as David Yonggi Cho’s harsher, more ventricular FVGs, superimposed over the adverbs as a form of personally experienced intensification, are related to the high-amplitude, ecstatic, near-glossolalic tongsonggido of Yoido Full Gospel Church’s Wednesday-morning healing service. Hence the nature and use of FVGs becomes an indexical icon not only of the voice quality of a speech register employed by a particular speaker, but also of the stylistic attributes of entire institutions—institutions that, given their size and their economic, political, and cultural reach, have influence over the everyday speech practices of Koreans.

Although I did occasionally hear Roh emit harsh FVGs on adverbs and adjectives, they were usually delivered comically, either to elicit laughter from his audience or in the performance of a comical allegorical figure (one also sees this sort of thing in the form of caricatures in films and on television). In general, there is a distinction between signaling the scalar intensity of some denotational content and signaling a focus on some denotational content in relation to other content. Roh’s whisper signaled a contrast within a conceptual framework rather than personal, i.e., speaker-centered, intensification. As a whole, his speech style favored this approach—whispered emphasis without FVGs—within a cultural framework of differently valorized forms of self-expression. The institutional cultures and liturgical styles of the two churches can be linked up with their respective prioritizations in church services of intensity of experience (the prototype being experiences of, and submission to, the Holy Spirit) and clarity of conceptualization (as a central feature of Christian intellectualism and leadership training), which, taken together, constitute a culturally significant opposition.

The point, however, is not that Pentecostals yell and Presbyterians whisper. Even if that were an accurate characterization, it would not be a surprising one (see Maltz 1985). Certainly there are Korean Presbyterian preachers who use plenty of harsh FVGs. Nor is the point merely that there is a syntactic shift from intensification to denotational focus—a distinction that can be fuzzy in sermons or any other explicitly illocutionary speech genre. The point is that Yoido Full Gospel Church and Somang Presbyterian Church stand as stylistic opposites within Korean Christian culture, which has consequences for speakers’ stylistic alignment to different models of Korean Christian personhood and their authorizing centers (see Agha 2005 on “enregistered voices” and Eckert 2008 on “persona style”). When the contrast is viewed ethnographically in relation to the communicative and social paradigms that intersect in the usage and ideological framing of FVGs more broadly, it becomes clear that there are competing frameworks for the sound shape’s appropriateness and
effectiveness of use. In the following section, I investigate an interdiscursively linked site in which related higher-order frameworks also intersect and are undergoing cultural change.

IV. A Ritual Center for the Expression of Personally Felt Intensity

Apart from segmenting and reproducing the prosodic FVG (see Section II above), most Korean speakers recognize the sound itself as an isolated utterance that is locatable in drinking events. Even when I would begin my questions with examples of the prosodic FVG, my informants often brought up the reactive FVG. They characterized it as a kamt’ansa (‘verbal reaction’) — a form of “self-talk” (Goffman 1978), like a “response cry” but in a distinctive affective mode — that people, especially men, make after taking a first sip of soju. While any Korean speaker might emit an isolated FVG when undergoing “traditionally” Korean sensory experiences, such as quaffing spicy Korean broth or sniffing fresh mountain air, in general, the exemplary emitter of the reactive FVG is the ajössi (see Agha 2003 on exemplary speakers). One conjures up an image of a noisy bar (sulchib) with loud, drunken men being served soju by women. As the men drink, they become increasingly animated and emotionally expressive, and their voices become increasingly “rough” (koch’ilda) or “husky” (hösük’i hada, swida), both in terms of overall voice quality and in their repeated use of both reactive and prosodic FVGs. In contrast, one would not expect to hear FVGs or husky voice quality at an American-style coffee shop or Italian restaurant. When one does hear such things, it is not uncommon to observe people rolling their eyes and mumbling “sikkiropta!” (‘noisy!’) under their breath.

The isolated FVG that emerges as a reaction to soju usually is produced as a voiceless aspirated velar plosive [k], followed by an FVG articulated on a high unrounded central [i] or back [u] vowel that resolves into an amplified whisper, as in: [kʰi:>::::]. Both informants and Korean colleagues have characterized this sound as a necessary ritual bracket for beginning the drinking event. Drinking events, in turn, create a social space in which men can, and are expected to, “release” not just stress or energy, but emotion as well (Lee 2007). These meetings can become so emotionally charged that experienced drinkers are fully aware that they could end up in either an embrace of friendship or a violent argument by the end of the evening. The FVG that follows the first sip of soju signals the entrance into this social space of emotional intensity.

So widespread and so well known is this sound that a large South Korean distiller, Jinro, recently launched an advertising campaign for a line of soju called Ch’amisul (‘true dew’) Fresh by creating a modified FVG to act as a sonic representation of the experience of, and reaction to, drinking the company’s soju. This campaign is based on what Jinro calls “voice marketing.” The company coins a sound that can be produced by people, that fits into the expressive repertoire described above, and that can be replicated to characterize not just the experience of drinking the liquor, but also the overall acoustemological (Feld 1996) experience that goes with it. The commercials and advertisements highlight the iconism of different sounds that one perceives during the event of drinking soju in Korea: the sizzling of meat over a flame (usually samgyopssal, a kind of Korean bacon), the sound of metal tearing apart as the cap is twisted from a fresh bottle, and the chorus of FVGs produced by friends sitting around a table.

Jinro’s FVG, however, is not the [kʰi:>::::] of the stereotypical ajössi, but instead a brighter, softer-sounding [kʰa:>::::] — the difference between the two being precisely the one that Yongch’an made clear to me on the street. Jinro even created a han’guł spelling for the sound: ㅛ. Figure 5 shows a Jinro poster depicting different people in their post-sip states (mouths open, tongues arched upwards) along with statements relating the reactive FVG to the experience of consuming the beverage. I have provided English glosses of the utterances superimposed over the images.
Although the drinking event depicted in the poster involves two men and two
two (as do the related television commercials), the campaign, like the soju-
soju-drinking event overall, remains male-centered. The women face directly into the
camera, as if addressing the potential consumer, while the men do not, as if serving
as an example with which the potential consumer should identify. Furthermore, the
men emphasize the sound produced by drinking soju, while the women emphasize
their own affective responses to listening to the sound.

A Google Image search for Jinro’s FVG reveals the term’s widespread use as a
descriptor for events of soju consumption even when Jinro soju is not featured.
Figure 6 features an image from an anonymous personal blog that displays a
picture of sizzling hagfish with a bottle of Doosan’s competing brand of soju,
Ch’ŏum ch’ŏrŏm (‘like the first time’), nearby. The blogger has written the following
caption beneath the image: 음장어 + 소주 = 카!!! (kkomjang’o + soju = k’ya!!! ‘inshore
hagfish + soju = [kʰjaʰː]!!!’). 28
Importantly, this meal is associated with maleness beyond the consumption of soju. With its long, quasi-cylindrical shape, the hagfish, like the eel, is often described as “good for men”—good for virility, for sexual drive, and for stamina (süt’aemina) (see Farquhar 1999, 2002 for similar examples in China). The combination of soju and a kind of male-gendered “medicinal meal” (Farquhar 2007) is understood to produce a verbal reaction in the form of an FVG.

We can begin to see how this isolable expression of intensity and its cultural license for expression are linked interdiscursively to, and, in a sense, mirror, the prosodic FVG and the legitimate intensification of speech. As a type of sound shape, the reactive FVG, which is associated with male-centered spaces of emotional expression, can be superimposed onto the phonological material of a vowel to deliver the prosodic intensification of speech. The prosodic FVG, which is associated with a speaker’s (not necessarily a man’s) claim to being the authoritative center of attention, invokes the reactive FVGs made by men during drinking events as they claim their right to express their emotions (as revealed when my informants made this connection). In both cases, the FVG suggests a speaker’s license to communicate personally experienced intensity. The sound shape links a communicative means of demonstrating and commanding expressive authority with a stereotypical locus and personification of expressive authority.

But things are changing for Korean women. Aside from participating in the drinking rituals that take place during obligatory military service for young Korean men, women have begun to take an increasingly prominent role in public drinking events in Korea (Kim and Kim 2008). One need only track the commercials for soju over the past twenty years to recognize a shift from a drinking experience popularly understood to be of and for males alone, with attributes of power and masculinity pervasive throughout, to one in which a specific model of the feminine becomes integral to the drinking context, even if still for a culturally accepted function of male enjoyment. In line with the soju industry’s widespread tactic over the past few years of lowering the alcohol content and marketing “softer” (tō pudurọpta) soju to women, it appears that the “hard” male sounds associated with the traditional site of male drinking are also being “softened” (cf. Bourdieu 1977, 1984[1979]:190–191 on la gueule and la bouche).

This shift toward a “softer” soju-drinking experience is emblematized by the shift in reactive FVGs. That is, the Jinro FVG iconizes the shift in Korean drinking culture in terms of both the reduced phenomenal intensity of the alcohol (softening the sound of the FVG) and the milder sensation in the throat (reducing the fricative intensity of the phonation). Indeed, this is what the director of research and development for one of the largest soju distillers confirmed for me in an interview in September 2009. He told me that the reactive FVGs that people make today are not the sounds that they made a long time ago (yennale) because the soju is not as strong and the “stimulation” (chagük) in one’s throat is not as intense. He began by explaining how the alcohol content was lowered from thirty to twenty-five percent, stabilized at twenty-five percent for a number of years, and then in just a few years dropped to less than twenty percent in connection with its increasing consumption by women. He told me, “As the level of drinking among women has increased, more than strong, more than hard [hadu˘], soft [sop’u˘t’u˘]—.”

Even if the Korean drinking event still remains basically male-gendered, there is a progressive shift toward authorizing women’s participation in it. Shortly following the statement above, the director explained how this lowering of alcohol content had resulted in a softening or “diluting” (hūtibakhada) of the alcohol and, by extension, of the reactive FVG. In his account, he produced the reactive FVG four times, both as demonstration and as reported speech; and in his characterization of the experience of drinking good, old-fashioned, strong soju, he described the feeling with the English loan-word “harsh” and superimposed a prosodic FVG over the [a] vowel in the initial syllable of this adjective. This is a very clear example of the relation between the reactive FVG, as a potential object of reported speech that expresses intensity, and the prosodic FVG, which comments on the intensity of whatever is being discussed—
captured here as an intradiscursive (Silverstein 2005) relation within the speaker’s explanation. In this case, what was being discussed was the intensely harsh feeling of drinking old-fashioned soju and the intensely harsh sound it produced in the drinker.

Excerpt 4: Softening both Soju and Sound

1. Actually, regarding stimulation, a long time ago...
   Chaguêk-e taehan sasil yennale-nûn
   
2. Nowadays, the Soju is not strong enough for the [kʰiːʃːː] sound to come out.
3. A long time ago, when the alcohol proof was high, when you drank, if the alcohol percentage was forty percent,
   Yennal-e alk’ol tosu-ka nop’ûl ttae-nûn masil tae-nûn alk’ol tosu-ka sasip to-ka toemyôn
4. also if you drank whisky, there is stimulation in the throat,
   wisûk’i-to masimyôn moggumông-uûro chaguêk-ûn,
5. because it stimulates, there is a “harsh” feeling, you know.
   A long time ago, when alcohol proof was high, when you drank, if alcohol percentage was forty percent,
6. Soju also had a “harsh” feeling, [kʰiːʃːː], and this kind of sound used to come out.
   Soju-ko h[al]sîhan nûggim-î ittjanayo,
7. “harsh” feeling, [kʰiːʃːː], and this kind of sound used to come out.
   Kûgô, oe han’guk sul chung-esô-nûn soju-ka kajang tokhan sulîöttgi taemmune,
8. Because soju is the very strongest of the Korean liquors
   tarûn sul-e pihamyon
9. if we compare [it] with other liquors
   tarûn sul-ûn t’akchu, moggolli-na, muô, iron kôs-ûn kûrôk’e tokhaji an’gi ttaemune sori-ka an-naonûnde.
10. because different liquors such as t’akchu or makkolli, um, are not strong, the sound doesn’t come out.
    Kûrôda ponikka sojuhamyon [kʰiːʃːː] han’n sorî-rûl naenûn kûrôhan, kû, súpsong-i mom-e paesô
    Come to think of it, if you have soju, the kind of, um, habit of making the [kʰiːʃːː] sound penetrates the body.}

11. and when you have soju it’s like you really associate [kʰiːʃːː] with it.
    Soju halttae-nûn [kʰiːʃːː]-rûl mani yônsgaranûn kôt kat’ayo.
Whereas the standard male reactive FVG used above is a response to the intensity of the sensation of taking a first sip of soju, the brightening and softening of the FVG in the Jinro advertisements and elsewhere corresponds to the lowering of the alcohol content, a less intense experience, and, for the Korean cultural context, a less male-centered event overall. Importantly, the lowering of the alcohol content is aimed not only at softening the experience of drinking, but also at reducing the level of drunkenness—or, at least, that is how it is often marketed. That is, both the weaker soju and the softer FVG are related to the increasing value placed on sobriety, especially among women, who are still considered relatively new to the events of public alcohol consumption. In the case of reactive FVGs, this shift from hard to soft, from drunkenness to sobriety, and from intensity to dilution parallels the co-occurring stylistic contrast between the ecstatic intensity of Yoido Full Gospel Church and the more sober restraint of Somang Presbyterian Church—a contrast made apparent by the use or suppression of prosodic FVGs in sermons.

Conclusion: FVGs and Paralinguistic Voicings

I have argued that the “illocutionary force” of FVGs in Korean prosody is linked to non-prosodic uses. My approach has been to demonstrate ethnographically the type-sourced interdiscursivity (Silverstein 2005) between prosodic FVGs, which are superimposed upon denotational content that a speaker invests with intensity, and reactive FVGs, as isolated utterances emitted in expressive reaction to an intense sensory experience. The reactive FVG is a well-recognized index of traditionally self-centered male behavior in which emotional expression is licensed. This self-centering function and its power to signal intensity of personal experience is replicated in the prosodic FVG without necessarily invoking maleness. Moving from the reactive FVG to the prosodic FVG, we can see how this particular voice gesture shifts from being the potential object of reported speech to being a prosodic form operating as a “sort of ‘commentary’ on the part of the speaker doing the reporting” (Lee 1997:278), i.e., from what Voloshinov called the “linear” to the “pictorial” styles of reporting speech (1973[1929]). At a more abstract semiotic level, the prosodic FVG can be seen as a kind of generalized indirect free style (Silverstein 1994, Lee 1997) of the reactive FVG, drawing on the authority of the stereotypical emitter of FVGs in the latter context—elder males situated at the center of the drinking event—for its performative capability to communicate prosodically both intensity and the legitimacy of its expression (regardless of whether or not a man is speaking). This indirect free style constitutes a kind of functional rank-shifting of the sound shape between the plane of utterance and the plane of prosody, where a sound shape that is reportable as an utterance actually becomes a paralinguistic feature in the framing of an utterance.

The linkage between prosodic and reactive FVGs brings together two higher-order paradigms of differentiation in one type of fricative sound: one most closely related to class differences, which can be seen in the contrast between two Christian denominations (Pentecostal and Presbyterian), and the other most closely related to the gendering of prototypical participants in public drinking events (male and female). As Jakobson and Waugh (1987[1979]:79) point out, speakers can become aware of such stylistic differences and map them along a time axis. In such cases, speakers experience different speech styles as modern or old-fashioned, with “time itself...enter[ing] into the verbal system as a semiotic value” (ibid.). In the case of FVGs, Korean speakers are becoming aware of these differences, with the harsher...
FVGs, both prosodic and reactive, seeming more and more to be a thing that “older” people do. This is due in large part to the real shifting social makeup of the ritual centers to which FVGs are related. Although the Yoido Full Gospel Church apparently continues to grow, its theology, starting in the late 1950s, was built upon a notion that the Holy Spirit would intervene in the lives and bodies of the suffering masses of postwar Korea. The Holy Spirit is said to appear as fire and sanitize the bodies of believers by burning out the sin, leading congregants to writhe ecstatically on the floor or speak in tongues (Harkness 2010a). Somang Church was founded in the late 1970s. Its message is that suffering and hardship are things of the past; its sermonic style and vocal standards—from whispered prayer to European-style classical singing—are seen as instantiations of a wealthy, healthy, modern, peaceful Korean present (Harkness 2010b). Meanwhile, although Korean soju continues to be consumed widely and in great amounts, it is viewed more and more as a traditional drink, for consumption primarily with Korean food in everyday bars and restaurants, especially when compared to the more recently imported luxury spirits, cocktails, and wines and their accompanying cuisines—none of which consistently gives rise to harsh reactive FVGs.31

With this in mind, we can now return to Sehyŏn’s story (Excerpt 1) and see the importance of producing an FVG on the first “kapchagi” but not the second. Sehyŏn would have learned FVGs from all sorts of sources such as television, bedtime stories, and the everyday speech of family members. Three women watch after Sehyŏn: his mother, his grandmother, and his aunt. I had heard his mother, the choir conductor, use prosodic FVGs on many occasions, but generally in a comical or markedly exaggerated way. More generally speaking, Sehyŏn would have accompanied all of these upper-middle-class women in their daily transitions between the loud, abrasive sounds of Seoul’s congested streets and markets and the quiet, softened soundscapes of Seoul’s coffee shops and Italian restaurants, where women of means congregate while their husbands work in offices or drink with colleagues at a local sulchip. Sehyŏn also would have experienced regular Presbyterian church rituals in which the description of a loving, caring God is delivered in a quieted, softened acoustic space. Angels, unlike king ants, are soft and peaceful.

We can also see that the broader shifting usage of the FVGs along a perceived time axis is immanent in the prosodic poetics of Sehyŏn’s narrative. The story is organized in terms of two “sudden” appearances, one of a villain and one of a hero. The villain introduces the personally felt intensity of terror; the hero introduces a calming moral order. The FVG marked the earlier phase of conflict and its absence marked the later phase of resolution. From chaos to comfort, the structure of Sehyŏn’s story is the same as the story that is told in Korean churches regarding Korea’s fifty-year rise from being one of the poorest, most war-torn countries in the world to being one of the richest and most developed. In these churches, it is often said that Christianity is responsible for producing a wealthy, peaceful, cosmopolitan South Korea.

And so, as Sehyŏn followed an allegorical formula of introducing a Christian hero to solve a problem of un-Christian behavior, he aligned a performance of different vocal styles with the different phases of the story. The combination of inclusion and erasure of the FVG serves as an emblem of this broader aesthetic-moral judgment. The phase of resolution here is told not through the loud, ecstatic expressions of Pentecostal Christians or men in a bar—that would much better approximate the phase of conflict—but rather, through the softer sounds of the more “modern” sites of consumption and of wealthy Presbyterian churches.

It appears that the negative-and-then-positive affect in the story is transferred via the FVG from the little ants and their feelings to the narrator’s empathetic voicing of their interests (Bakhtin 1981:275–336; Hill 1995; Silverstein 1993; 1994:50–53, 59n11; 2000). In the story, the ants themselves didn’t utter the FVG. In fact, they didn’t utter anything at all. Rather, it was the narrator who uttered the FVG when describing their
situation. In Sehyôn’s story, the perspective of a silent character within the narrated event “breaks through” as the perspective of the narrator himself via meaningful phonological adjustment within the event of narration. This is consistent with the way the narrator explicitly aligns himself with different characters of the story—a fact made clear by two of Sehyôn’s other gestures, namely, his facial expression of fear at the arrival of the king ant (as the perspective of the little ants) and his kinesic performance of the angel’s efforts to drive the king ant away (as the perspective of the angel).

Fricative Voice Gestures, then, do not merely act as shibboleths of a particular identity or social category (although they certainly can do so). Nor are FVGs simply part of the acoustic environments of loud bars or rowdy churches (although they certainly contribute). Rather, the speaker’s deployment of the FVG cues the speaker’s alignment to, and stance on, different frameworks of authority and social differentiation invoked in different kinds of communicative events. To use an FVG is not necessarily to perform a specific character or type—be it man or woman, elder or youth, monster or angel, ecstatic Pentecostal or sober Presbyterian. Although the inhabiting of such identifiable roles can be the ultimate indexical entailment of FVG usage, the FVG does not always produce these particular entailments. I have presented the examples above to show that the FVG actually is a multi-purpose emblem within intersecting and parallel paradigms of differentiation in a rapidly changing society. It is a feature of various expressive registers of authorized intensity in South Korea that members of younger generations, the educated upper classes, and a particular stratum of Protestant Christianity increasingly characterize as a remnant of past—and passé—cultural forms.

Notes

Acknowledgments. A shorter version of this paper was presented at the Annual Meeting of the American Anthropological Association as part of a session titled “The ends of prosody: Ways of speaking as ideological means and ends,” December 4, 2009. Thanks are due to the organizers, Mark Sicoli and Matthew Wolfgram, as well as to Robin Queen for her helpful comments as the session’s discussant. Donnald Brenneis, Hisun Kim, Jae-eun Kim, Kiho Kim, Kwangkun Lee, Seayong Park, and Michael Silverstein read and commented on subsequent drafts. Special thanks are due to Paul Manning, Miyako Inoue, Paul Garrett, and the anonymous reviewers, whose comments greatly improved the paper. The research, writing, and revision process for this article was funded by the Mellon Foundation/Hannah Holborn Gray Fellowship; numerous FLAS Title VI Scholarships for Korean; and a University of Chicago Center for East Asian Studies Travel Grant. Any errors are my own.

1. I have followed the McCune-Reischauer system for the Romanization of Korean, with one minor exception: I have hyphenated post-positional particles instead of inserting a space.
2. See Footnote 7 below for an explanation of the transcription.
3. In commercials in the U.S., Europe, Japan, China, and elsewhere, one sees similar such things in the exaggerated audible exhalations of consumers after a first drink of beer.
4. Cf. Goffman’s (1979) notion of the role-partial of a “principal,” i.e., having a stake in the denotational content.
5. Crucially, though, this is not the sound of spitting. One informant corrected my pronunciation of the FVG by telling me not to make too much of a “saliva sound.”
6. The reader can try an exercise in order to see how this works. Put your articulators in place as if you were going to pronounce a voiceless velar fricative, as in “Bach.” Now treat the voiceless velar fricative as a kind of sustained drone, and try to enunciate a phrase over it without phonating with the vocal cords. It will sound a bit like static, but you should be able to hear your phrase. It is the same principle as a whisper, only the vocal tract is far more obstructed. This is essentially a much more complex version of the FVG, which, as a gesture, simply takes one vowel (or diphthong) in a phrase and replaces vocal-cord phonation with a supralaryngeal site of frication. Although phonation traditionally has been described as a laryngeal setting (Ladefoged 1975, Laver 2009[1980]), we might say that vowel phonation for prosodic FVGs shifts from the vocal cords to some supralaryngeal site of frication.
7. The term “fricative vowel” has been suggested for vowels that, through coarticulation with consonants, adopt into a phonated vowel some articulatory feature of a fricative (Ladefoged and Maddieson 1996). This can be represented with a diacritic for voice quality and a subscript symbol to indicate fricative articulation combined with a vowel. For example, [a], with the under-ring diacritic and a subscript [x], would represent a voiceless velar FVG as a fricative [a] vowel.

However, in my observation, FVGs are not necessarily produced as an allophonic addition through coarticulation (cf. Standard Chinese and Czech in ibid.). Although a velar FVG might proceed from, and seem to have some coarticulation with, a velar consonant (e.g., Excerpt 1), coarticulation with a velar consonant is not necessarily a condition for a velar FVG (e.g., Excerpt 2). Until the phonetic relationship of FVGs to the category of fricative vowels more generally is further clarified, my approach to notating FVGs is simply to add a superscript [x] to the vowel symbol to indicate consonant-like frication plus vocalic articulation without phonation of the vocal cords. All of the FVGs presented in this paper are produced with velar frication (hence my choice of [x]), but, in my observation, the FVG also can be produced with other loci of friction and could be notated accordingly. For example, an [a] vowel with an aryepiglottal FVG would be [a¹]; with a pharyngeal FVG would be [a³]; with a uvular FVG would be [a²]; and so on, according to sites of friction.

Furthermore, FVGs must be distinguished from the so-called “emphatic” consonants of Arabic (see Watson 2007). FVGs are prosodic features superimposed onto the distinctive features of a Korean vowel, whereas the Arabic emphatic consonants form a whole phonological class defined by distinctive features. While FVGs are “emphatic” in the sense of emphasizing some stretch of speech, i.e., having some relationship to reference and predication, phonological “emphasis” in Arabic refers to the spread of articulatory correlates (e.g., pharyngealization) from the “emphatic” consonant to adjacent vowels.

8. Ko (2002) observed that the locus of the emphatic lengthening for scalar adverbs in Korean is the last segment of the first syllable: a vowel for open syllables, or any coda consonant for closed. The first syllable, “chong,” is closed, so the velar nasal [ŋ] provides the phonological material for lengthening.


10. One strand of research on gesture has focused primarily on the role of visible gestures in relation to speech (see Kita 2001, Braddick 2009). Building on the notion of “paralanguage” introduced by George Trager (1958, 1961), this area was a major focus of Ray Birdwhistell in the 1950s and 1960s, culminating in his 1970 collection of essays, Kinesics and Context: Essays on Body Motion Communication—a volume replete with close analysis of verbal-kinesic coordination and the development of “kinegraphs,” a sort of kinesic equivalent of the IPA. The study of what is called “non-verbal communication” has been at the heart of research by psychologists of interaction such as Starkey Duncan and Adam Kendon. Kendon follows from the fact that visible gesture is coordinated with—is metricalized by—utterance to argue that gesture can be used as utterance itself (2004). It becomes a pantomime of the totality and thus acts as an index, pars pro toto. Similarly, David McNeill’s (2005) term “catchment” describes the poetic chunking of visible gesture in discourse (see also Kataoka 2009). See also Handler (2009:283–4) on taken-for-granted notions of the “naturalness” of visible gesture.

11. In phonology, the term “gesture” usually comprises three sub-gestures: categorial (relatively more vocalic or consonantal), articular (place features), and initiatory (airstream and glottal state) (Lass 1984:282). This understanding of gesture was adopted and elaborated among proponents of what is generally called the “Motor Theory of Speech Perception,” which claims that “phonetic information is perceived . . . [as] the intended gestures of the speaker that are the basis for phonetic categories” (Liberman & Mattingly 1985:1). The muscular “gestures” of phonation are, cognitively, “mirrored” in the decoding of them by a perceiver. This last perspective is strikingly similar to George Herbert Mead’s (1934:46) insofar as communicative interaction is reduced to a process in which speakers respond to one another with “like gestures” in the emergent creation of sociality and mutually agreed-upon meanings.

12. “La richesse phonétique du gazouillis cède la place à une restriction phonologique.” Presumably, Jakobson’s use of the term “sound gesture” follows from Jespersen’s (1918; qtd in Jakobson and Waugh 1987[1979]:186) notion of sound gestures “nesting in vocabulary,” an example of which can be seen in Jespersen (1922:136): “Most children learn to say ‘no’ before
they can say ‘yes’ . . . Sometimes the n is heard without a vowel: it is only the gesture of ‘turning up one’s nose’ made audible.”

13. According to Bolinger (1946), audible gesture may be phonemic, intonational, or syntactical. Bolinger’s speculations on the genesis of any given meaningful gesture prior to its conventional usage in discourse (e.g., the “I don’t know” act in 1986:198–9, 211) were based on the assumption that “both intonation and gesture, like Antaeus, cannot survive without the earth. In other words, we have a mixed system, expressive at base but with adaptations that differ from culture to culture” (1986:198). What Bolinger meant by the word “expressive,” he explained, was Saussurean non-arbitrariness. We can see, then, that behind Bolinger’s empirical, rigorous, detailed, and compelling descriptions of the paralinguistics of speech lurks a theory of language evolution that is concerned with the social process of moving from natural, unmediated expression to socialized, conventional speech. Bolinger made this view explicit in 1978, when he famously referred to intonation as “a half-tamed savage” (475). He wrote, “to understand the tame or linguistically harnessed half of him, one has to make friends with the wild half” (ibid., qtd. in Queen 2001). Taken to the extreme, such speculations easily approach the “‘Ding-Dong,’ ‘Pooh-Pooh,’ or ‘Bow Wow’ theories of yore . . . cast in the idiom of language origin” (Silverstein 1994:40)—or, for that matter, the “Yo-Heave-Ho” theory (also known as the “Yo-He-Ho” theory).

14. This suggests some socially stratified limits of awareness, specifically in the realms of segmentability and metapragmatic transparency (Silverstein 1981). We can compare the use of FVGs with Irvine’s (1990) example of noble and griot speech registers in a rural Wolof community, where generalizing assumptions about the ability of members of the “lower” griot social caste to control emotional expression intersect with local understandings of the situational license, indeed, expectation, of nearly all members of the community to be “griotlike” and on occasion to use the emotionally expressive speech register. In Korean speech, however, the explicitly labeled and grammatically and lexically coded speech levels do not relate directly to affectivity in speech, nor are they understood to share the qualities of any particular social categories (see Gal and Irvine [1995:974] on iconization and Gal [2005:35n5] on rhematization). Furthermore, the FVG can be used in any of the speech levels, as the examples throughout the paper show. Finally, the perceived permissibility of using the FVG as an element of prosodic form depends on the speaker’s position as, in a sense, privileged and therefore justified in commanding an interlocutor’s attention. It is when a speaker does not realize that he or she does not have this social permission that the FVG is considered rude (as when someone uses a speech level that is deemed inappropriate to the communicative event). Interestingly, however, the traditional Korean image of the Confucian upper-class person, the yangban, closely approximates Irvine’s description of the self-controlled Wolof noble. See also Ekman (1979–79) on “display rules” for the expression of emotion and intensity; Caffi and Janney (1994) and Besnier (1990) for a collection of references; and Agha (2005) for more on registers.


16. We can interpret the congregation’s “Amen!” as a kind of large-scale version of what Victor Yngve (1970) originally called “back-channeling.”

17. We can compare the use of FVGs with Irvine’s (1990) example of noble and griot speech registers in a rural Wolof community, where generalizing assumptions about the ability of members of the “lower” griot social caste to control emotional expression intersect with local understandings of the situational license, indeed, expectation, of nearly all members of the community to be “griotlike” and on occasion to use the emotionally expressive speech register. In Korean speech, however, the explicitly labeled and grammatically and lexically coded speech levels do not relate directly to affectivity in speech, nor are they understood to share the qualities of any particular social categories (see Gal and Irvine [1995:974] on iconization and Gal [2005:35n5] on rhematization). Furthermore, the FVG can be used in any of the speech levels, as the examples throughout the paper show. Finally, the perceived permissibility of using the FVG as an element of prosodic form depends on the speaker’s position as, in a sense, privileged and therefore justified in commanding an interlocutor’s attention. It is when a speaker does not realize that he or she does not have this social permission that the FVG is considered rude (as when someone uses a speech level that is deemed inappropriate to the communicative event). Interestingly, however, the traditional Korean image of the Confucian upper-class person, the yangban, closely approximates Irvine’s description of the self-controlled Wolof noble. See also Ekman (1979–79) on “display rules” for the expression of emotion and intensity; Caffi and Janney (1994) and Besnier (1990) for a collection of references; and Agha (2005) for more on registers.

18. When Cho retired a few years ago, Lee Young-hoon, who had been the head pastor at the Full Gospel satellite church in Los Angeles, took over as head pastor. Lee and the other pastors at Yoido Full Gospel Church follow Cho’s sermonic style; their sermons are marked with numerous FVGs, even during stretches of relatively quiet, restrained speech.

19. In a personal communication, Kiho Kim pointed out to me that this is a common strategy among Korean politicians, especially when they give speeches in large public spaces.

20. There are exceptions. In one of his sermons, Somang Church’s head pastor Kim Chich’ol uttered the Sino-Korean compound p’yōngsaeng (‘lifetime’) with a harsh FVG on the ㅇ of the first syllable. Classified as a noun, the term actually includes an adjective, “ordinary, flat, even” (p’yōng, ပြါ, 朴) and a noun, “life” (saeng, საქ, 佛山). Just as if he had been using a separate adjective, he placed the FVG on the Sino-Korean word p’yōng so that the pronunciation became p’y[ʊŋ]sāngsaeng. In so doing, he was able to add intensity while using the
Sino-Korean lexicon, which is associated with a speech register higher and more refined than the native Korean stratum of vocabulary (Cho 2002).


25. In a personal communication, Saeyoung Park pointed out that something like “voice marketing” has been used in the naming of Korean brands of noodles such as Hururuk Kkusu. Hururuk is a denotationally iconic adverb meaning ‘slurpingly’ or ‘with the sound of slurping,’ and kkusu means ‘noodles.’ Although it is not uncommon to hear a person produce a bright, breathy FVG on an [a] vowel after slurping down some spicy Korean ramen, the emphasis in the descriptive name hururuk is on the sound of the consumption of the noodles rather than the sound of one’s reaction to it.

26. The shift from [i] to [a] also signals a shift from the category of “dark” vowels to “bright” vowels in Korean’s system of vowel harmony. Such vocalic ablaut has semantic and pragmatic implications in Korean’s rich inventory of denotationally iconic words (Martin 1962, Kim-Renaud 1976, Sohn 1999, Lee and Ramsey 2000), one of which is an augmentative-diminutive distinction.

27. Note how the command is to “react” or “exclaim” (kamt’anhada). My informants described isolated FVGs as expressive “reactions” (kam’tansa, 감탄사, from the Chinese 感歎詞 “feeling-sigh-word”).


30. The phrase can also mean that one ‘gets accustomed to’ something.

31. I was privileged to spend an evening with Korean colleagues who indulged me by trying a number of different alcoholic beverages, from whisky to wine, in order to see whether any of them might produce an FVG like that of soju. The consensus was that none of them did.

References

Agha, Asif

Agha, Asif

Austin, John

Bakhtin, Mikhail M.

Bauman, Richard

Besnier, Niko

Birdwhistell, Ray L.

Bolinger, Dwight
1946 Thoughts on “Yep” and “Nope.” American Speech 21(2):90–95.

Bolinger, Dwight
1961 Generality, Gradience, and the All-or-None. The Hague: Mouton de Gruyter.

Bolinger, Dwight

Bolinger, Dwight

Bolinger, Dwight
Bourdieu, Pierre  

Bourdieu, Pierre  

Braddick, Michael, ed.  

Brenneis, Donald  

Brown, Roger and Albert Gilman  

Caffi, Claudia and Richard W. Janney  

Cho, Young-Mee Yu.  

Eckert, Penelope  

Ekman, Paul  

Ervin-Tripp  

Farquhar, Judith  

Farquhar, Judith  

Farquhar, Judith  

Feld, Steven  

Fonagy, Ivan  

Gal, Susan  

Gal, Susan and Judith Irvine  

Goffman, Erving  

Goffman, Erving  

Grammont, Maurice  

Gumperz, John J.  

Handler, Richard  
Harkness, Nicholas

Harkness, Nicholas

Hastings, Adi and Paul Manning

Hill, Jane H.

International Bible Society

Irvine, Judith T.

Jakobson, Roman

Jakobson, Roman, Fant, C. Gunnar M.; , and Halle, Morris

Jakobson, Roman

Jakobson, Roman

Jakobson, Roman and Linda Waugh

Jespersen, Otto

Jespersen, Otto

Kataoka, Kuniyoshi

Kendon, Adam

Kim, Wooksoo and Sungjae Kim
2008 Women’s Alcohol Use and Alcoholism in Korea. Substance Use & Misuse 43:1078–1087.

Kim-Renaud, Young-Key

Kita, Sotaro

Ko, Eon-Suk

Korean Bible Society

Labov, William
Labov, William  

Ladefoged, Peter  

Ladefoged, Peter and Ian Maddieson  

Lass, Roger  

Laver, John  

Lee, Benjamin  

Lee, Iksop and Robert Ramsey  

Lee, Sang-Gil  

Liberman, Alvin and Ignatius Mattingly  

Maltz, Daniel  

Manning, Paul  

Martin, Samuel  

McNeill, David  

McQuown, Norman  

Mead, George Herbert  

Newman, Stanley  

Osburg, John  

Queen, Robin  

Sapir, Edward  

Sapir, Edward  

Silverstein, Michael  

Silverstein, Michael  
Silverstein, Michael

Silverstein, Michael

Silverstein, Michael

Silverstein, Michael

Silverstein, Michael

Silverstein, Michael

Silverstein, Michael

Sohn, Ho-min

Trager, George L.

Trager, George L.

Trudgill, Peter

Voloshinov, Valentin. N.

Watson, Janet

Yngve, Victor

Department of Anthropology
Harvard University
William James Hall
33 Kirkland Street
Cambridge, MA 02138
harkness@fas.harvard.edu