

## Placing social types through prosodic variation: An investigation of spatial meanings in Mainland China

Robert Xu\*

**Abstract.** This study examines how prosodic features evoke the spacial aspects of interactional meanings of well-known social types in Mainland China. Prosodic features (duration, pitch, voice quality) of the scripted performances of 18 prominent social types in China were measured acoustically and grouped by cluster analyses. Commonalities among types within each group were identified through a detailed analysis of meta-linguistic commentary collected from the internet. This paper focuses on three meaningful clusters: powerful bureaucratic types, disembodied voices, and “in-your-face” types. Members of each cluster share prosodic combinations and social profiles. More importantly, character types within each cluster index a specific interactional locale. Appropriation of their associated features could reproduce the social dynamics that is typical in that locale. The results highlight the situated use of sociolinguistic variables, and show that the prosodic features pattern structurally in the performances while indexing the historical-spatial settings of social interactions. This paper also considers place as an interactional and relational product of meaning making by these prosodic features.

**Keywords.** character type; prosody; place; social meaning; stylistic performance; Mandarin Chinese

**1. Introduction.** Prominent social types such as “Valley Girl” and “Bro” in America easily evoke social images that are associated with specific linguistic styles. These character types, or what Agha (2003, 2005, 2011) calls characterological figures, are widely conventionalized social abstractions enregistered with linguistic features. They function as landmarks in a broad social-semiotic landscape (Eckert 2016, in press). Their distinctive stylistic features provide the speech community with materials for the construction of personae. They are dialogical in nature, defined in large part by the kinds of interactions they engage in. Therefore, the performance of a character type, once abstracted in the social-semiotic landscape, becomes a sign to typify not only the social types, but also the time and place where the interactions occur (Agha 2005).

This study investigates the spacial aspect of the interactions evoked by a constellation of character types in Mainland China. The few sociolinguistic studies on social types in Chinese communities mostly focus on their historical aspect. Zhang studies the emergence of Cosmopolitan Mandarin through social types such as Yuppies working in international companies in Beijing (Zhang 2005, 2017) and Lifestyle TV Hosts in Tianjin (Zhang 2017). Cosmopolitan Mandarin contrasts sharply against the northern and “standard”-sounding Beijing Mandarin, with features such as rhotacization and r-lenition carried by the character type “Smooth Operators” (Zhang, 2005, 2008) and features such as fronted dentals carried by the character type “Alley Saunterer” (Zhang 2005). In another series of studies in Hong Kong, Kang and Chen trace a novel gender stereotype, Kong girl. This derogative stereotype of a materialistic, vain, attention-seeking young girl came into being in the context of the shifting heterosexual marketplace thanks to increasing cross-border marriage, with rising anxiety in both male and female roles and values

---

\* I would like to thank Patrick Callier for his early contribution to this project. Author: Robert Xu, Stanford University ([robxu@stanford.edu](mailto:robxu@stanford.edu)).

(Kang and Chen 2014, Chen and Kang 2015). While these studies focus on the historical aspect of social types, the spatial dimension of the types is implicitly and inevitably mentioned, such as the international company the “yuppies” work in, the old Beijing alleys, and the changing border between Hong Kong and Mainland. However, this dimension has become secondary to the historical dimension.

This study focuses on the spacial dimension of the character types and theorizes space as the immediate context of the interaction, what Giddens (1984) termed “locale”. Schegloff (1972) identifies two types of place-formations: the “geographical or “absolute” terms for “national, regional municipal locations”, and the “relational” terms for the immediate situations where interactions take place. Sociolinguists have generally focused on place as geographically defined region and community. What has often been ignored is the spatial elements at the micro-level in the interaction – the world social actors experience as a direct context of the interaction (Johnstone 2004). Goffman (1959) uses “setting” to refer to all the physical surroundings and items that facilitate an actor’s performance. Giddens (1984) calls the use of space to provide the setting of interaction “locale”, a concept essential to contextualizing the interaction. A locale is centered around and mediated by the body, as seen from the perspective of the actor. As the body is a living being, a locale is inherently temporal.

To investigate locale evoked by character types, I choose post-lexical prosodic features - such as intonation, voice quality, and speech rhythm - as the primary linguistic resources to investigate. Prosody, in contrast to segmental features, ties closer to the performativity of the “inner self” (Eckert 2017). They deliver the first impressions of the physicality of the speaker, as the critical ingredients for the speaker’s identity (Podesva and Callier 2015). Furthermore, locale is closely related to and bounded by the body, and the body is bridged to language through voice (Laver 1980). Many social meanings associated with locale manifest through highly malleable prosody (Podesva and Callier 2015).

This study investigates how stylistic prosodic features pattern structurally to character types of similar social behaviors in Mainland China, and the spacial meanings in the form of interactional locale encoded in the character types’ performances. I examine scripted performances of 18 character types by 10 speakers in the lab, and use cluster analyses to link acoustic measurements of their speech with the qualitative analysis of social comments from the internet. These comments highlight the perceptions of these character types, and reveal a wide range of social meanings about interactional place in the Chinese social landscape, evoked by prosodic performances.

**2. Methods.** This section reports the design, collection, and analysis of the lab-based study to attain the acoustic data for prosodic features in the character types’ performances.

**2.1. DATA COLLECTION.** The selection of the character types aims to represent a wide range of well-known types in Mainland China. They are all subjects of constant social comments, in the forms of discussions, parodies, memes, and imitations. No participant in this study reported any difficulty recognizing any of them. Most of the chosen character types are as enregistered as “Valley Girl” in the U.S. The final selection is diverse in gender, age, social status, and region. More than 50 character types were initially considered, but the list was narrowed to 18 so that the character types in the final list were verbally performable, visually identifiable, distinct to each other, and not offensive.

For each type, a paragraph-long preparatory script was written in the “voice” of the character. Some images from the Internet that accurately represent the specific character types were chosen to normalize individual variation in interpreting the types. Three target sentences (1) – (3)

were designed. They are character type neutral, so that every character type could utter them naturally. They are also designed to incorporate many potentially socially meaningful variables in Mandarin Chinese, such as nasality, rhotacization (*erhua*), narrow focus and intonation type, for this and future studies.

- (1) zhège shì shuíde? [tʂɤ4 kɤ5 ʂi4 ʂei2 tɤ5]  
This-CL be whose  
'Whose is this?'
- (2) nǎlǐ mǎi de huā? [na3 li3 mai3 tɤ xua1]  
Where buy POSS flower  
'Where did you buy the flowers?'
- (3) wǒ huítóu gěi nǐ dǎdiànhuà. [wɔə3 xui2 tʰoʊ2 kəi3 ni3 ta3 tɤən4 xua4]  
I later give you make-phone  
'I will call you later.'

Thirteen native Mandarin speakers (9F/4M, aged 23-29) participated in the recording in a sound-treated booth at the Chinese University of Hong Kong. All of them studied or worked at the university. Three of them were not analyzed either because they articulated difficulty in performing the types due to their ability to act or confidence level, or because the recording was too loud for some character types. This left ten speakers (7F/3M) analyzed. The speakers come from different areas in China, and may speak another Chinese language or dialect. All of them use Mandarin Chinese as a primary language of communication.

Each speaker was presented with one character type at a time. Each character type slide included the character type label, the preparatory script, images of the type, and the target sentences. The speakers were instructed to emulate the character type's voice. After practicing as many times as needed, they performed the script, and then read the target sentences as a continuation of the performance. They spoke directly into a Sony PCM-D50 recorder with external microphones. Recording sessions took 40–60 minutes depending on the speaker.

**2.2. DATA ANALYSIS.** Three groups of prosodic features were measured: Pitch variability (F0 mean and F0 range over a sentence), duration variation (syllable duration and pairwise vocalic variability index (VnPVI)), and voice quality variation (H1\*-H2\* and cepstral peak prominence). In addition, vowel formants were measured just in case some syllables in critical prosodic positions show interesting vowel quality.

The sounds were aligned to the script using the FAVE forced alignment system (Rosenfelder et al. 2011) to obtain phone- and syllable-level time alignments for each character type rendition, with the Chinese syllables treated as new entries in the dictionary. The alignments were further hand-corrected at both the syllable and phone level.

The acoustic measurements were conducted in Praat (Boersma and Weenink 2015) using a script that extracted measurements every 10 ms across the length of the stimulus. F1-F3 were estimated using Praat's formant tracker, supplied with parameters adjusted for each speaker. F0 was measured using the autocorrelation method and pitch minima and maxima adjusted for each speaker. H1 and H2 were estimated by picking maxima in a narrowband FFT spectrum in a window around F0 and 2×F0, respectively. We calculated H1\* and H2\* - H1 and H2 corrected for influence of proximal vocal tract resonances. Cepstral peak prominence was extracted from a cepstrum by searching for peaks within the speaker-specific pitch range window. Vocalic nPVI was measured as the difference in vowel durations in pairs of adjacent syllables, normalized by the mean duration of the vowels in each pair.

These measurements can be categorized into three groups of prosodic features in the analysis. First, for the durational patterns, the vocalic nPVI measures the variability of vowels in the utterances. A positive cluster of VnPVI suggests less syllable timing. The duration of three syllables of the first target sentence, representing different boundaries of the intonational phrase (mid-boundary *ge* and final boundary *de*) and the locus of the natural focus (*shui*), were also measured, as well as the F1 of the final neutral tone of the utterance. Second, the pitch pattern is represented by the mean F0 level, pitch variability in the form of F0 standard deviation, and F0 range between the 10% and 90% points of the utterances. Third, the phonation pattern is measured through H1-H2 for creakiness-breathy and smoothed cepstral peak prominence (CPP) for the regularity of the vocal folds vibration.

Since the purpose of the analysis is to observe the patterning of the prosodic features across character types, rather than conducting hypothetical testing about the performances, cluster analysis is used, with the help of “cluster” “mclust” and “pvcust” packages in R. The goal of the cluster analysis is to combine similar prosodic features and similar character types into groups, to observe commonalities in social characteristics that unify clusters of character types exhibiting similar prosodic patterns. Two levels of cluster analyses were conducted. The variable-level analysis sorted the character types into three groups ( $k=3$ , positive, neutral, negative, indicated by “<<”, “--”, “>>” respectively in following figures) according to the similarities of each variable. The type-level analysis grouped all the character types into an un-predetermined number of clusters considering the overall similarities across all variables. Five clusters of character types emerged from this type-level analysis.

**3. Results and discussion.** The results of the variable- and type-level cluster analyses show meaningful patterning. If the speakers did not draw on prosodic features to portray meaningful character types, this would not be expected. Five clusters of character types result from the type-level analysis. They seem to be, in most cases, meaningful groups. In other words, within most of the five groups, the character types share many prosodic features, and also behave similarly in some social aspects. In the following sections, I will highlight three groups by connecting their shared prosodic features and the common perceived social meanings of their styles.

**3.1. DURATIONAL PATTERN AND THE BUREAUCRATIC STYLE.** Two character types - *lingdao* “Bureaucrat Leader” and *laoganbu* “Retired Party Cadre” - are grouped together by their shared prosodic features (Table 1). Notably, they both have higher VnPVI than the other types. Mandarin Chinese has been reported to be a syllable-timing language (Mok 2011), with minor variability expected across vowels in an utterance. These two character types, however, display a less syllable-timing speech rhythm, with their VnPVI indicating that they have the highest variability of vowel durations across the utterances.

Clustered character types	Duration				Pitch			Phonation	
	final	focus	<i>ge</i>	VnPVI	level	St.Dev	range	H1-H2	CPP
<i>lingdao</i> “Bureaucrat”	--	--	>>	>>	<<	--	--	--	--
<i>laoganbu</i> “Retired Cadre”	--	--	>>	>>	<<	--	--	--	>>

Table 1: Clustering of prosodic features for the bureaucratic types

A closer look at the data shows that the vocalic variability mainly results from the special intonational phrasing, performed across speakers (seven speakers for Retired Cadre and five speakers for Bureaucrat Leader). Figure 1 shows the spectrograms and pitch contours of

utterance (1) for the Retired Cadre performed by a speaker. Two interesting patterns can be observed. First, the speaker divides the five-syllable utterance into two separate intonational phrases, evident in the pitch resetting at the third syllable. Second, the second syllable, even though it supposedly has the unstressed neutral tone, was followed by a pause and/or lengthening to mark the end of the first intonational phrase. This observation is also consistent with the results of the duration measurement of the *ge* syllable (Table 1). That syllable is much longer in the performances for these two character types than those of other types. In addition, the lengthened *ge* was often produced with neither its full tone (a falling tone) nor a neutral tone (with a low pitch following the preceding Tone 4), but rather a long mid-level pitch contour, not among the canonical tone categories in Mandarin Chinese.

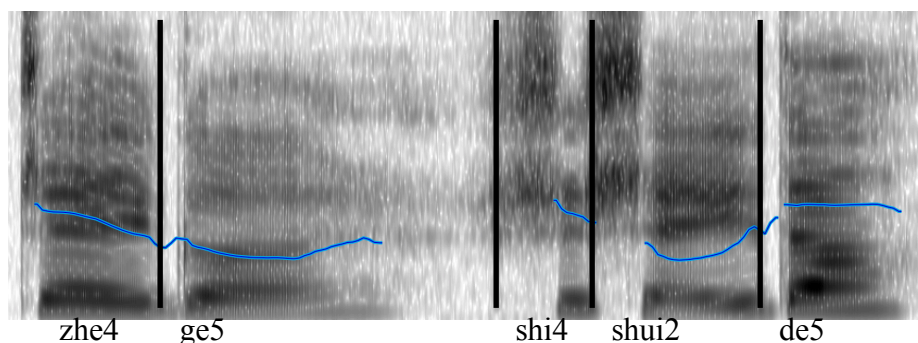


Figure 1. Spectrogram and pitch contour for laoganbu “Retired Cadre”

I argue that this special durational pattern is performed to reflect the stereotype of bureaucrats, widely known as *guanqiang* “bureaucrat tongue” across Chinese communities. *Guanqiang* is frequently discussed in public discourse in connection with its use in regular tedious meetings hosted by the bureaucrats. Their speech, often without much important or useful information, is full of time-fillers like the lengthened *ge* syllable and the pause that follows. The speech rate slowed down by the pause and the lengthening helps to deliver the message in a clear manner with fewer errors. It also makes it easier to place informational focus. This style is further associated with the uselessness and the condescension of such speech and with the people who produce it. In the largest Chinese question-and-answer on-line community *Zhihu* ([zhihu.com](http://zhihu.com), a counterpart of America’s Quora), where people also often share their life experiences, many ask and talk about the style. One of the users asked about his experience with the stereotype and whether there was a way to avoid the style.

*“I found the bureaucrat giving speeches in meetings the way people working for the government would do. The elderly boss always tries to use some bullshit to fill in the time....Why can’t the state-owned company hold effective and brief meetings as in international companies...?”*

The style is rooted in the centralized power dynamics in party-led China. The Retired Cadre character type is part of this power structure; while the Bureaucrat lives in local governments, state-owned companies, and the education system that share a mirroring power structure. The style is so imbued with power dynamics that it becomes cultural capital to maintain power. People new to this system have to acquire this style in order to become a part of it. In this sense, the style dialogically not only connects the speaker and the audience in a public

speech genre; but more importantly, the performance engages in the creation of the power structure that contrasts the powerful and the subordinate.

Furthermore, this marked intonational pattern can be appropriated in stylistic performances to evoke the dialogical power dynamics outside of political and workplace discourse. The phrase *da-guanqiang* “use the bureaucrat tongue” and its synonym *da-haha* “use the fake laugh” (“haha” refers to the fake laugh the bureaucrats often use in their interactions with subordinates) were used to depict verbal behaviors that treat others’ problems perfunctorily, in interactions characterized by an asymmetric relationship in power or resource. Crucially, the special intonational phrasing with pause and lengthening is an important component of the behavior that brings the specific power dynamic to life in mundane interactions.

The special durational pattern facilitates the creation of the power structure discursively and iconically. Pauses and lengthening could create anticipation of increasing information or uncertainty (Goldman-Eisler 1958), bring tentativeness or emphasis to speech (Brown and Levinson 1978), or indicate that the speaker is searching for the most delicate way to express the upcoming content (Diamond 1996). While these features add to the ritual and ambiguous performance, they also help to create an image of authority, by iconically linking to the use of time in the meetings. Time can be seen as a currency of power in this context. In the setting of the institutional meetings, those in power, rather than those with actual information, take the stage first and talk longer. Only the person in power can afford to use (referentially) “meaningless” lengthening and pauses to take up more time.

From the analysis above, it becomes clear that the durational pattern in the performances of the Bureaucrat and the Party Cadre helps to establish both the temporal and spatial profiles of the interaction. As historical products, the bureaucrat style is inherently old-fashioned, representing the “traditional” social dynamics in communist China. Equally important is the locale that is typified by this style: meetings inside the party or state-owned corporations. The bureaucrats that take on this style hold the power in the room and speak towards an audience that is silent and subordinate in the genre of a public speech. Their speech does not license any interruption even though it is filled with pauses and fillers. The bureaucratic character types in this study live in the same chronotope (Bakhtin 1981) as state-owned corporation workers in Zhang (2005). Connecting to Zhang’s study, the locale of the bureaucrats - the state-owned corporation - both provides a space that resists social change and allows the bureaucrats and their silent audience to stay behind the change - a change that has brought an open forum, free discussion and an awakening of individuality.

3.2. DISEMBODIED BREATHY VOICE. This section focuses on a cluster of character types that is associated with breathy voice. Figure 2 shows the phonation features measured for all the character types. Smoothed cepstral peak prominence (CPP, y axis) identifies how periodic the vocal folds are vibrating (the higher, the more periodic); while H1\*–H2\* depicts whether the voice is more creaky or pressed (negative) or breathy (positive). A clustering group (boyinyuan “broadcaster”, *xiaozi* “petit bourgeois”, *peiyin* “voiceover actor”, Table 2) takes the right bottom corner, indicating that they share a breathy voice.

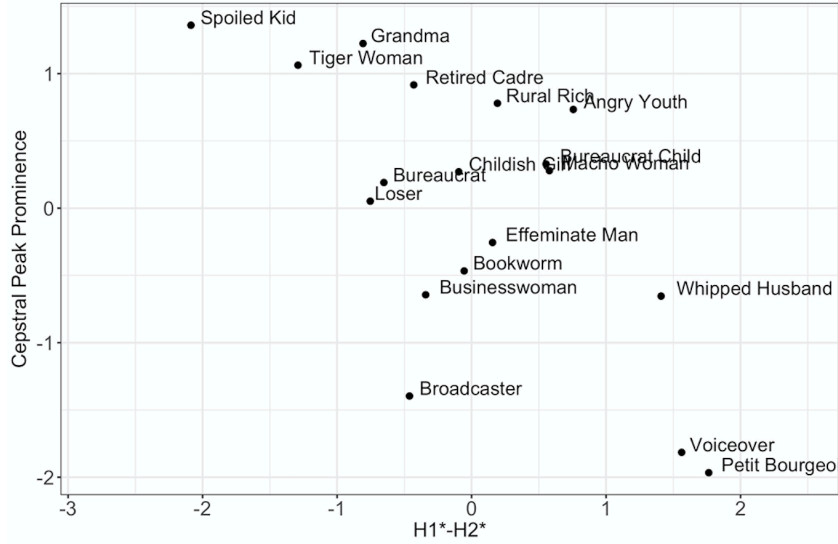


Figure 2 Normalized voice quality measurements across character types

Clustered character types	Duration				Pitch			Phonation	
	final	focus	ge	VnPVI	level	St.Dev	range	H1-H2	CPP
<i>xiaozi</i> “bourgeoisie”	--	--	<<	<<	<<	--	--	>>	<<
<i>peiyin</i> “voiceover”	--	--	<<	--	<<	<<	<<	>>	<<
<i>boyinyuan</i> “broadcaster”	<<	--	--	--	<<	<<	<<	--	<<

Table 2: Clustering of prosodic features for the disembodied-voice types

Meta-discursive analysis shows that these character types with a breathy voice and low F0 are de-individualized voices. For example, the normativity and standardness of the News Broadcaster eliminates any marks of regional features as well as individuality, making it a pan-China voice. Often commented as the broadcasting style on social media and characterized as being “normative, solemn, instigating, modern, proper, affable” (S. Zhang 2003), the style becomes a voice of the institution over that of an individual. Even Jia and Li (2013), a textbook on producing the broadcasting style for college students majoring in broadcasting, recognizes that the broadcasting style might be interpreted negatively for sounding “pretentious, coming from above, even dehumanizing”.

The voice performances in dubbed western movies are also recognized as a distinguished style known as *peiyin-qiang* “dubbing tongue”. While the style has a lexical element - uncommon phrases and words from translation and uncanonical lexical tones - its voice quality plays an essential role. Much like the broadcasting voice, the dubbing voice is regarded as inauthentic for an everyday Chinese person. It is a transbodied performance in nature, indexing non-Chineseness, and is mocked on the internet as sounding “unnatural, pretentious, superficial and frivolous” (quoted from *Zhihu*).

The Petit Bourgeois is a newly emerged young urban white-collar type, marked by their taste, desire, and a cosmopolitan lifestyle. Starbucks emerges as a cultural landmark for this group. It creates an image of working at leisure, by quantifying life by time and money spent in a café. Moreover, Starbucks contrasts sharply with traditional places of social interaction in China,

thereby creating a new social category: the young westernized urban middle class. They are also a product of social change, as noted in a cultural blog by Qi Zhang (2012).

“The motivation for this culture is the forceful presence of capitalism and bureaucracy, under which the petit bourgeois ... could only entertain themselves in a masochistic way. ... restricted by their economic foundation, they could not really reach the goal. Wandering between hope and despair give birth to the mocking attitude, and the petit bourgeois culture.”

The Starbucks life is not just a state of fact but rather a luxury in desire. Therefore, the voice of petit bourgeois is a disembodied vehicle for a consumerist lifestyle that is dislocated from the context of everyday Chinese interaction.

In summary, with the use of the breathy voice to iconically index a degree of refinement that is spate from the messiness of everyday life, the news broadcaster, the voiceover actor, and the bourgeois are voices transcending their body, and live in a public space and a new social category that separates them from ordinary people and run-of-the-mill interactions.

**3.3. ICONIC FEATURES OF THE HEIGHTENED PRESENCE.** This section will focus on a group of character types that share many prosodic features, and all appear to be a highly affective, aroused presence in interactions. As shown in Table 3, they share high F0 level and variability, some of them also have attention-grabbing lengthening at on-focus or utterance final syllables, as well as a pressed, shouting-sounding voice quality (top left corner of Figure 2).

Clustered character types	Duration				Pitch			Phonation	
	final	focus	ge	VnPVI	level	St.Dev	range	H1-H2	CPP
<i>xionghaizi</i> “Spoiled Child”	>>	--	<<	<<	>>	>>	>>	<<	>>
<i>niangpao</i> “Effeminate Man”	--	>>	<<	--	>>	>>	>>	--	--
<i>dia</i> “Childish Girl”	>>	>>	<<	<<	>>	>>	>>	--	--
<i>pofu</i> “Tiger Woman”	<<	<<	<<	--	>>	>>	>>	<<	>>
<i>dama</i> “Neighborhood Grandma”	--	--	<<	<<	>>	--	--	--	>>

Table 3: Clustering of prosodic features for the “in-your-face” types

These types also share similar social positioning and behaviors. They come to life in highly intimate or affective interactions, often in a loud, in-your-face manner. Specifically, many of those character types demands a heightened presence. This presence allows them to navigate the private social space, where they could poke around, fish for information, and even manipulate their interactants, like the nosy neighborhood grandma and unreasonable tiger woman. By using high F0 in level and variability, coupled with the “shouting” voice, they are able to perform in-your-face, intimidating acts that help them achieve their communicative goals.

Furthermore, these character types are gender specific. All of them happen to be women, children or effeminate men. High F0 level and variability may help to evoke the stereotype that women and children speak with a higher pitched voice and tend to be more expressive. However, not every female character type belongs to this group, e.g. *nvhanzi* “Macho Woman” and *nvqiangren* “Strong Businesswoman”. Even though both types have the word *nv* “female” in them, they did not stimulate performances using a high F0 level stereotypical of female speech. Instead, this group gathers a specific kind of female type. While Neighborhood Grandma and Tiger Woman are comprehensibly loud and agitated females, I will now elaborate on the other



three types - Spoiled Kid, Effeminate Man, and Childish Girl, which form a sub-cluster of their own within this cluster in the type-level analysis.

Spoiled Kid and Childish Girl (Girls who are *dia*) share both prosodic profiles and social behaviors. Both interact with others to gain something: Spoiled Kid asks from adults, Childish Girl asks from men. The act of Childish Girl to show their *dia*-ness is called *fadia* (literal: deliver the *feminine quality*) or *sajiao* (literal: scatter the feminine charm), and is common across East Asia (See Moon (under review) for *aegyo* in Korea, Starr (2015) for sweet voice in Japan, and Yueh (2012) for *sajiao* in Taiwan). Yueh (2012) points out that *sajia* is an act where a girl talks or behaves like a child for persuasive purposes, by embracing the image of helpless and weakness, taking advantage of the attention and care that men could offer. In order to achieve this goal, the Childish Girl borrows the speech features of a Spoiled Kid to sound weak but demanding. The goal of *sajiao* is not just to show sweetness and innocence, but to manipulate men to do what a Childish Girl wants. As a tag line from the movie “Women Who *Sajiao* Have the Best Fortune” says: “Dating is never just treating others a dinner; it is a battle of a woman overturning another woman.” Analyzing a similar social behavior *aegyo* in Korea, Moon (under review) points out that although women use *aegyo* as an interactive weapon to dominate men, the act reinforces the gender order by positioning women on the side of pleading. She also argues that *aegyo* is a designed performance depending on the linguistic execution and proper interactive context. Childish Girl cannot sound like a child just any time; they have to execute this performance with an interactive goal and in certain immediate interactive contexts.

Similar to the Childish Girl, *niangpao* “Effeminate Man” (literal: Sissy Cannon) uses similar F0 features and final lengthening. This character type is associated with a conventionalized style called *niangniang-qiang* “girly tongue” that is often received negatively. Table 4 displays some advice from *Zhihu* on how not to sound girly, which requires attention to both pragmatics (Row 1) and voice quality (Row 2). A stereotypical Effeminate Man gossips, judges, uses clean and polite language, his voice thin and high, often using features from Southern dialects to sound soft and delicate. Although the Effeminate Man is not perceived as demanding and calculated like the Childish Girl, they are similar in the overwhelmingly feminine quality that could be traced back to child speech through a recursive semiotic process (Irvine and Gal 2000).

1	Avoid using Cantonese accent. Don't be mean to women. Try not to show your disdain when judging. Use some foul language. Be more direct and rude. Avoid gossiping with girls.
2	Practice your voice to make it thicker. Don't speak from your throat, try to find the heavy low voice. If your voice condition is not perfect, speak less and do more.

Table 4 Internet advice for guys to avoid sounding effeminate

The gendered repertoire of character types in this group fits the stereotype of their social behavior and reveals the locale of these character types. The in-your-face gossiping and manipulation are believed to be women's territory. An old Chinese saying goes, “*nan zhu wai, nv zhu nei*” (men deal with business outside of the family, women deal with business within the family). There is a dichotomy of social domains between men and women. That is why men are counseled to “gossip less” in order not to sound effeminate (Table 4, Row 1). In fact, someone who gossips a lot (especially for a man) is described as *popo-mama* (granny-and-mommy sounding); and gossip itself is termed *jiachang-liduan* (big and small business within the family). It is not uncommon to see real-life male counterparts of the Neighborhood Grandma but their linguistic performance is not enregistered. This dichotomy explains why these character types live in a private space within the family or a close neighborhood – the locale of the interactions they typify

is “within the family”. The linguistic features that they use iconically link to their interactive behavior. The lengthened focus and final syllables, together with high F0 mean and variability, allow them to conveniently grab the attention of their interactants in an expressive manner. The high and variable voice pitch helps them to become highly affective, sometimes aroused or intimidating. The high voice pitch and the shouting voice form a style that shows up as a heightened presence in a small interpersonal distance, as the loudness resembles the closeness of the bodies. It is within this intimate space that these character types live and interact.

**4. Conclusion.** While sociolinguistic interviews provide large amounts of naturalistic speech, the performances in this study highlight salient features of the situated use of variables, as well as stylistic resources that are unlikely to emerge in interviews. The results of the cluster analyses based on the acoustic measurements were interpreted with meta-linguistic analysis. Three meaningful groups of character types emerged, each with distinct prosodic patterns and social profiles. The powerful bureaucrat types (Bureaucratic Leader, Retired Party Cadre) employ marked international phrasings with pause, mid-sentence lengthening, and pitch resetting. The disembodied voices (News Broadcaster, Voiceover Actor, Petit Bourgeois) are performed with breathy voice and low pitch level and variability. The “in-your-face” types (Neighborhood Grandma, Childish Girl, Spoiled Kid, Effeminate Young Man, Tiger Woman) use high F0 level and variability, final/focus lengthening, and a loud, shouting-like voice quality. As the meta-linguistic analysis shows, to understand the social meanings carried by these prosodic features, it is vital to look at not only what kind of person these features index, but also what kind of interaction is typified. This includes all the elements of the “voicing structure” (Agha 2013) – the message, the speaker and the other interactants - as well as the temporal-spatial context of the interaction where the message becomes relevant and that all the interactants inhabit.

The temporal-spatial context of interaction is typified by the character types through their use of prosodic features, in many cases in an iconic manner. For example, the pause and lengthening of the bureaucrats illustrate power taking the form of protracted time in public meetings, where the subordinate audience is silenced. The carefully controlled breathy voice indexes refinement and performance dislocated from the self. The attention-grabbing high F0, shouting voice, and focus/final lengthening mimic the over-the-top interactional style of the in-your-face types resemble attention-grabbing interactional style in a private setting. Through the cluster analyses, where clusters of prosodic features correspond to meaningful clusters of character types, it becomes clear that prosodic features are important resources the participants draw on to perform these types. These features take on social meanings by patterning structurally and iconically to evoke the spatio-historical settings of interactions.

This spatio-historical setting of an interaction is what Giddens (1984) terms “locale”. Locale as the immediate context of interactions is temporal, bodily, and meaningful. As discussed above, locale is closely related to modes of interaction such as speech situation and the genre of interaction. For instance, a contrast between the public sphere and private space is shown between groups with breathy voice and shouting voice. Arousal interaction in private spaces can be inferred from the combination of shouting voice and high voice pitch. Social power dynamics between the speaker and the audience can also be inferred from durational patterns, as seen in the analysis of the bureaucrat style. The character types may also typify specific physical settings of interaction such as meetings in state-owned companies, newsroom, and Starbucks café.

Furthermore, the enregistered prosodic features can be appropriated in stylistic practice to evoke a type of corresponding interactive situation and the dialogical power dynamics that come

with it. As a result, the social meaning can be adopted out of context even when the interaction no longer takes place in its specific locale. As an additional example, the news broadcaster style can evoke an impression of officialness and unnaturalness, sounding fawning and inauthentic. This suggests that the social meanings of these prosodic features are so closely tied to a locale that they evoke the specific relational dynamics in new interactions.

The results also suggest that social change is materialized through the place evoked by the character types. The bureaucrats stay with the communist party politics, the disembodied voices live in new social categories, the “in-your-face” types engage in private interactions that mimic traditional family and neighborhood relationships. The intrinsic connectedness of the time-space texture invites further discussion of Bakhtin’s (1981) choronotope, in producing cultural values from sociolinguistic variables to a broad social-semiotic landscape. By forming chronotopes through the use of sociolinguistic variables, the character types both contrast with each other (private space vs. public space; riding social change vs. stay behind social change) and relate to each other (the clustering of types) in place and time; and therefore construct a social-semiotic landscape. In the performances in this study, a social-semiotic landscape of the rarely studied modern Chinese Mainland materializes through the alignment of prosodic performances with interactional meanings. The participants, all lay speakers without acting training, demonstrate the ability to present this landscape through performances with the enregistered linguistic features.

## References

- Agha, Asif. 2003. The social life of cultural value. *Language and Communication* 23. [https://doi.org/10.1016/S0271-5309\(03\)00012-0](https://doi.org/10.1016/S0271-5309(03)00012-0).
- Agha, Asif. 2005. Voice, footing, enregisterment. *Journal of Linguistic Anthropology* 15 (1). 38–59. <https://10.1525/jlin.2005.15.1.38>.
- Agha, Asif. 2011. Large and small scale forms of personhood. *Language and Communication* 31 (3). 171–80. <https://10.1016/j.langcom.2011.02.006>.
- Bakhtin, M. M. 1981. Forms of time and of the chonotope in the novel: Notes toward a historical poetics. In Michael Holquist (ed.), *The dialogic imagination: Four essays*. 84–258. Austin, Texas: University of Texas Press.
- Boersma, Paul & David Weenink. 2015. Praat: Doing phonetics by computer.
- Brown, Penelope, and Stephen C. Levinson. 1978. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Chen, Katherine H.Y. and M Agnes Kang. 2015. Demeanor indexicals, interpretive discourses and the ‘Kong Girl’ stereotype: Constructing gender ideologies in social media. *The Journal of Language and Sexuality* 4(2): 193–222.
- Diamond, Julie. 1996. *Status and power in verbal interaction: A study of discourse in a close-knit social network*. Amsterdam: John Benjamins.
- Eckert, Penelope. To appear. The individual in the semiotic landscape.
- Eckert, Penelope. 2016. Variation, meaning, and social change. In Nikolas Coupland (ed.), *Sociolinguistics: Theoretical debates*. Cambridge University Press.
- Eckert, Penelope. 2017. The expressive side of language. Talk at LSA Linguistics Institute.
- Giddens, Anthony. 1984. *The constitution of society: Outline of the theory of structuration*. Berkeley and Los Angeles: University of California Press.
- Goffman, Erving. 1959. *The presentation of self in everyday life*. New York: Doubleday.
- Goldman-Eisler, Frieda. 1958. The predictability of words in context and the length of pauses in speech. *Language and Speech* 1(3). 226–31. <https://10.1177/002383095800100308>.

- Irvine, Judith T. & Susan Gal. 2000. Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of language: Ideologies, politics, and identities*. 35–84. Santa Fe: School of American Research Press.
- Jia, Ning, and Hao Li. 2013. 论‘播音腔’的含义 On the meaning of the ‘broadcasting style.’ *Jin Chuanmei (Communication Today)* 8.
- Johnstone, Barbara. 2004. Place, globalization, and linguistic variation. In Carmen Fought (ed.), *Sociolinguistic variation: Critical reflections*. 65–83. New York: Oxford University Press.
- Kang, M. Agnes & Katherine H.Y. Chen. 2014. Stancetaking and the Hong Kong Girl in a shifting heterosexual marketplace. *Discourse & Society* 25(2). 205–20. <https://10.1177/0957926513515587>.
- Laver, J. D. M. 1980. *The phonetic description of voice quality*. Cambridge, UK: Cambridge University Press.
- Mok, Peggy Pik-ki. 2011. The acquisition of speech rhythm by three-year-old bilingual and monolingual children: Cantonese and English. *Bilingualism: Language and Cognition* 14(04). 458–72. <https://10.1017/S1366728910000453>.
- Moon, Kyuwon. Under review. Authenticating the fake: Linguistic resources of *aegyi* and its media assessments.
- Podesva, Robert J. & Patrick Callier. 2015. Voice quality and identity. *Annual Review of Applied Linguistics* 35. 173–94. <https://10.1017/S0267190514000270>.
- Rosenfelder, Ingrid, Joe Fruehwald, Keelan Evanini & Jiahong Yuan. 2011. FAVE forced alignment and vowel extraction program suite. <http://fave.ling.upenn.edu>.
- Schegloff, Emmanuel A. 1972. Notes on a conversational practice: formulating place. In Pier Paolo Giglioli (ed.), *Language and social context*. 95–135. London & New York: Penguin.
- Starr, Rebecca L. 2015. Sweet voice: The role of voice quality in a Japanese feminine style. *Language in Society* 44 (1). 1–34. <https://10.1017/S0047404514000724>.
- Yueh, Hsin-I. 2012. *The tactic of the weak: A critical analysis of feminine persuasion in Taiwan*. Iowa City: University of Iowa dissertation.
- Zhang, Qi. 2012. 小资和小资文化 Petite bourgeois and petit bourgeois culture.
- Zhang, Qing. 2005. A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society* 34(03). 431–66. <https://10.1017/S0047404505050153>.
- Zhang, Qing. 2005. 2008. Rhotacization and the ‘Beijing smooth operator’: The social meaning of a linguistic variable. *Journal of Sociolinguistics* 12(2). 201–22.
- Zhang, Qing. 2005. 2017. *Language and social change in China*. New York: Routledge.
- Zhang, Song. 2003. *中国播音学 Chinese broadcasting announcing*. Communication University of China CUC Publishing House.