Speech Composition of Food Review Influencers

Jeannette Shijie Ma
Project paper for the course "Research Seminar Social Technology"
Media Technology MSc program, Leiden University, The Netherlands
www.mediatechnology.leiden.edu
July 2019

***Abstract* — Food review influencers play an important role in shaping consumer behaviors. This type of large-scale synchronization relies on a successful word composition of the speech. In order to find out whether the influencers introduce more sensory profile in their speech (as modern food packages do), this study used python and manual classification to perform a text mining on the subtitles of two popular food review YouTube channels. The result shows the influencers used more common ingredient and food names and less sensorial descriptors in their speech. This might suggest that our brains learn new items better by comparison and relations instead of profiling.**

I. INTRODUCTION

Social communications and activities have been shaping our brain during human evolution. Language was one of the social tools we used to effectively communicate our perceptions (1). Communicating about food while eating together enhanced our synchronized memories and social bonds. There is an extreme case of food communication – the food review influencers. Social media influencer has been regarded as a key to viral marketing of products (2). Online communication makes an impact on our empathy and trust on the related communities (3). As more and more food varieties are accessible, consumers have less time to make a choice by themselves. Those influencers as food reviewers plays an important role in directing consumer choices and preferences.

Audiences' information involvement is an important aspect leading to the success of a YouTube influencer (4). For food review influencers, it is important that their speeches achieve connections with their audiences so that the audiences understand what they could expect when eating the food they saw on the screen. Therefore, it is especially interesting to analyze the speech composition they used for successful large-scale synchronization.

Sensory profile has been used in food communication with consumers (e.g. on some modern packages). They are shown to affect consumer choices (5). However, sensory profile and perception descriptors were mostly invented by experts and less commonly used in daily life conversations. It is not known whether those profiles are the best way to communicate food perceptions. The speech composition of popular food review influencer might provide us with an overview of possibilities in food communication.

This study used text mining approach to answer the research question: Does food reviewers use more sensory profile in their speech composition for food communication?

II. METHOD

Step 1: subject selection

In this study, two popular YouTube channels were selected as subjects based on their subscription number and main topics. The first channel (1.5 million subscribers) reviews fast and takeaway food (hamburger, pizza etc.). Every video focus on one food product. The food was individually reviewed and filmed in a quiet place (e.g. at home). No other people appear in the videos and the influencer only talk to the audience. Thus, no communication happens in the videos, the speeches used are only controlled and contributed by the influencer himself during filming. The second channel (4.6 million subscribers) reviews street and restaurant food. Every video focus on one journey or place, so multiple food could be involved in a single video. Since the influencer interact with the local people or restaurant owners during his journey, their conversations were also in the subtitles.

Step 2: subtitle sample scraping and word count

After content checking and removal of non-food related videos, the first 50 most popular videos from each channel was selected (100 videos in total). Their subtitles were scraped by an external website: savesubs.com into txt. format. For each channel, Python with the Panda package was used to transform uppercase to lowercase and generate a word count result into an excel sheet.

Step 3: manual text cleaning and word classification

From the excel sheet, non-food related words were removed if they are not in one of the pre-decided categories: 1. cooking method, 2. Ingredient & food names, 3. sensorial perceptions, 4. Appearance. Each remaining word was classified into one of these four categories. The sum of word count in every category is divided by the total word count (after text cleaning) to get the "category usage frequency" for each word category.

Step 4: sub-categories

For the "ingredient & food name" category, words were further classified into "ingredients" (e.g. 'garlic') and "food product" (e.g. 'burger'). For the "sensorial perceptions" category, words were further classified into "general perception" (e.g. 'smell') and "descriptors" (e.g. 'sweet'). For the sub-category "descriptors", words were further classified into "tastes" (e.g. 'sweet'), "flavors" (e.g. 'fresh') and "textures" (e.g. 'crunchy').

III. RESULTS & DISCUSSION

Pie charts were generated with the results from each channel subject. The comparison was made for all main categories, the sub-categories from "ingredient & food names" and "descriptors".

Figure 1 shows the main category usage frequency of the two channel subjects. "Ingredients and food names" was the mostly used word category in their speeches for both channnels (~65%) followed by "Sensorial perceptions" (~20%). For these main categories, both street food and takeaway food influencers are similar in speech compositions.

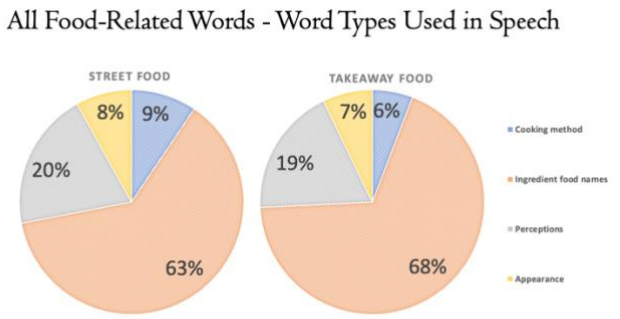

Fig. 1. Composition of four main categories

Figure 2 shows the sub-category results from the main category "ingredient and food names". For the street food influencer, ingredient names were much more used than food product names. However, for the takeaway food influencer, food product name was more used than ingredient names. In street food communication, freshness and the ingredients used might be much more frequently discussed or introduced by the influencer. In the takeaway food case, ingredients might be already visible or less interesting to mention. The comparison between common food products seems to be more important, which might have boosted the usage of food product names.

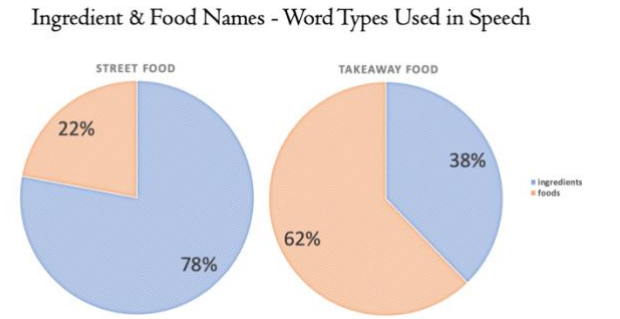

Fig. 2. Composition of sub-categories from 'ingredient & food names'

As shown in figure 3, both the street food and takeaway food influencers had a quite balanced usage on "taste", "flavor" and "texture" words. Although much more varieties of texture descriptors were used by both influencers, the total word usage frequencies of these three perceptions were not very different.

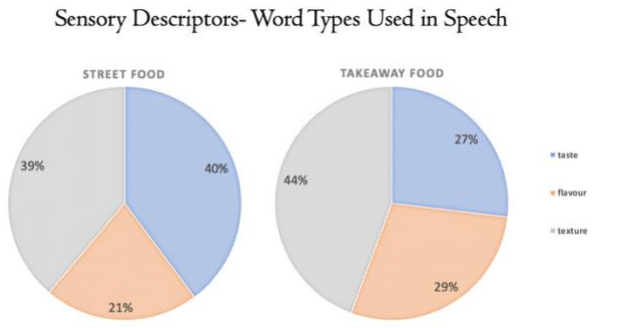

Fig. 3. Composition of sub-categories from 'descriptors'

Apart from the main result, there were other observations during the manual word cleaning process: The street food influencer used many location names (e.g. 'Indian') and environmental descriptions (market, walking etc.) to specify local culture and destination of the journey. Price related words (dollar, deal, worth etc.) were frequently used in both influencers' speeches.

In professional sensory evaluations, descriptors or made-up descriptors were mainly used to communicate a food product. For example, words like "cheesy" and "creamy" would be more frequently used to describe a cheese flavored cake instead of directly naming the possible ingredients. However, in daily conversations, descriptors seem less used when communicating foods. Palatability (e.g. 'this is yummy!') and ingredients (e.g. 'it tastes like pork') are more frequently used and easier to understand. The results show that the food reviewers use less sensory profile in their speech composition for food communication than food and ingredient names. When the influencers are trying to evoke synchronization of large-scale audiences, their success might rely on the fact that they belong to the audience group and share the same "language of food". Instead of sensorial descriptors, more frequent usage of common ingredient and food names might be more efficient in communicating a food to public audience.

Different from robots, human have better ability to learn new concepts from single or few examples (6). The pattern recognition ability of human is a well-developed brain function throughout the evolution course. Similar to the quest for computers to recognize a chair from an office picture, it is difficult for machines to identify a "pizza pepperoni" flavor. However, for human it might be easier relating "pizza pepperoni" to a similar flavored potato chip product. Food communication would be much more difficult if people found words like "salty", "smoky" instead of "pepperoni flavor" on a chip package.

It is reasonable for sensory evaluation experts to discuss in professional languages in order to correlate with machine analysis results. However, in terms of large-scale food communication (food package, advertisement etc.), synchronization might be more efficiently achieved by evoking relative memories than absolute memories. Instead of making effort to translate sensorial glossaries into consumer languages, it might be better to introduce a more natural discussion task to professional sensory panel. Relating a new product to common ingredients or other popular products could be a better choice in both consumer tests and expert sensorial evaluations. This could achieve more transparent food communication and a better match of new products to their target consumer groups.

IV. FURTHER RESEARCH AND RELEVANCE

This study only focused on 100 videos of two influencers. Although this selection controlled topic and style consistency, the result might be biased by the speech routine of the same influencer. For further researches, it would be interesting to do the same analysis on videos from multiple influencers. Also, comparing the results with YouTube comment text mining on the same videos would be beneficial to understand how the synchronization works.

Extending the discussion into a broader view, we may confirm the theory that social categorization makes intergroup relations more defined and clearer (7). When introduce a new item or individual to a social network, human might be naturally

built to compare the item or individual to the other items or individuals they knew and find the most similar ones to form a "stereotype". These stereotypes could be used later to efficiently get knowledge of another new item or individual (8). As a new individual, it is also easier to learn complex networks by relations and categories. Instead of profiling, comparison, relation and classification might be a more natural and efficient way of social cognition and communication.

REFERENCES

[1] Dunbar, R., Gamble, C., & Gowlett, J. (2014). Thinking big: How the evolution of social life shaped the human mind. Thames & Hudson.

[2] Mohr, I. (2014). Going viral: An analysis of YouTube videos. Journal of Marketing Development and Competitiveness, 8(3), 43.

[3] Preece, J. (2004). Etiquette, empathy and trust in communities of practice: Stepping-stones to social capital. J. UCS, 10(3), 294-302.

[4] Xiao, M., Wang, R., & Chan-Olmsted, S. (2018). Factors affecting YouTube influencer marketing credibility: a heuristic-systematic model. Journal of Media Business Studies, 15(3), 188-213.

[5] Carrillo, E., Varela, P., & Fiszman, S. (2012). Effects of food package information and sensory characteristics on the perception of healthiness and the acceptability of enriched biscuits. Food Research International, 48(1), 209-216.

[6] Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. Science, 350(6266), 1332-1338.

[7] Crisp, R. J., & Hewstone, M. (2007). Multiple social categorization. Advances in experimental social psychology, 39, 163-254.

[8] Howard, J. W., & Rothbart, M. (1980). Social categorization and memory for in-group and out-group behavior. Journal of Personality and Social Psychology, 38(2), 301.