Research Collection Lee Kong Chian School Of Business

Lee Kong Chian School of Business

2-2023

Boosting persuasion: The attention benefits of multiple narrating voices

Hannah H. CHANG
Singapore Management University, hannahchang@smu.edu.sg

Anirban MUKHERJEE

Amitava CHATTOPADHYAY

Citation

MARKETING

Boosting Persuasion: The Attention Benefits of Multiple Narrating Voices

by Hannah H. Chang, Anirban Mukherjee, and Amitava Chattopadhyay

Image Credit | Windows

The "Voice Numerosity Effect": When hearing different voices narrating a marketing video facilitates persuasion.

☑ INSIGHT | FRONTIER   06 Feb 2023

"An iPod…a phone…and an internet communicator." With these words, Steve Jobs introduced the iPhone in his 2007 keynote address.

Firms such as Samsung, Tesla, and Google introduce new products in tightly-scripted keynotes. These presentations, delivered at marquee events such as Mobile World Congress (MWC) and Google I/O, are central to shaping consumer and investor perceptions of new products. Successful keynotes such as Apple's introduction of the Mac, the iPhone, and FaceTime, have become etched in our collective memories and created value for Apple shareholders. For instance, Apple's introduction of the iPhone at Macworld 2007 was associated with an approximately 7% increase in stock value. In contrast, widely panned keynotes such as Qualcomm's keynote in Consumer Electronics Show (CES) 2013, failed not only in persuading investors of the innovativeness of firms' new product pipeline but in driving sales.

RELATED CMR ARTICLES

"Ad Lib: When Customers Create the Ad" by Pierre Berthon, Leyland Pitt, & Colin Campbell. (Vol. 50/4) 2008.

"De-Marketing Obesity" by Brian Wansink & Mike Huckabee. (Vol. 47/4) 2005.

A crucial component of product introduction keynotes is the product introduction video that visually depicts a product whilst one or multiple off-screen narrators enumerate its features and benefits. Such videos are first presented at keynotes, then played on the news, on social-media websites such as YouTube, and also used by technology reviewers and influencers such as MKBHD. The production technique is known as voice-over narration and it has been used in marketing videos for decades in wildly popular advertisements such as Billy Crudup's voice-over in Mastercard's 1997 priceless campaign that originated the tagline, "there are some things that money can't buy, for everything else there is Mastercard."

In a forthcoming paper in the *Journal of Marketing Research*, we investigate the role of voice (narrator) numerosity in marketing videos (Chang, Mukherjee, and Chattopadhyay 2022). For example, consider the following two real-life examples: a product video introducing Apple's AirPods Max had two narrating voices while a product video introducing Apple's new Macbook Pro had a single narrating voice. Does the difference in the number of narrating voices influence consumers' attention and subsequent behaviour?

Despite the ubiquity of the product introduction format, and more broadly, the ubiquity of video advertising across different media and formats from television in the 70s and 80s to online video advertising today, the role of voice numerosity as a potential strategic design element to enhance the persuasiveness of marketing videos had not yet been examined. Importantly, research in sound and cognition has found that amongst many elements, the human voice plays a unique role in human cognition—prior research in neuro science and psychology shows that the human brain has evolved to be intricately sensitive to the human voice such that it is recognised by babies before they understand language, activates unique regions in the brain, quickly draws attention, and evokes immediate and greater processing (Horowitz 2012; Belin, Fecteau, and Bédard 2004; Rutten et al. 2019). Furthermore, prior research suggests that some specific audial features of the human voice (e.g., volume) may affect how listeners' perceive a narrator (Apple et al. 1979). For example, voice pitch may influence consumers' perception of narrator competence (Chattopadhyay et al. 2003). These findings highlight a paradox: while the use of such features (e.g., using a lower pitch voice actor) can increase persuasion by influencing consumers' positive perceptions of the narrator, they may also distract from and reduce consumers' attention to the product message that is the point of the marketing video (Chaiken and Eagly 1983; Grewal, Gupta, and Hamilton 2021).

In line with this conceptualization, we examined the implications of voice numerosity in marketing communication videos such as product introduction videos and video advertisements on consumer behaviour. The research involved the careful and detailed analysis of very large-scale, real-world datasets on crowdfunding and video advertising through the use of machine learning and sophisticated econometric models, as well as the use of carefully planned and designed laboratory experiments to isolate psychological mechanisms.

Figure 1: Visual Summary

THE ATTENTIONAL BENEFITS OF VOICE NUMEROSITY 🔊

MORE VOICES PERSUADE

Chang, H., Mukherjee, A., & Chattopadhyay, A. (2022). More voices persuade: The attentional benefits of voice numerosity. *Journal of Marketing Research*. https://doi.org/10.1177/00222437221134115

By analysing multiple datasets* using machine learning and text mining techniques, we found:

The persuasive power of a marketing video is enhanced when a video employs more narrators, particularly when it conveys a simpler message

Voice numerosity occurs because the use of more (different) narrators helps draw consumers' attention to the spoken message

Voice numerosity exerts consistent and economically significant effects on a wide array of downstream behaviours including purchase

*Including a dataset with 11,000+ product videos and over 3.6 million transactions in Kickstarter, a leading crowdfunding platform

WHAT IS AT WORK? THE "VOICE NUMEROSITY EFFECT"

Consumers tend to be more persuaded, when they hear a message narrated by different voices in succession in a video than when they hear the same message narrated by a single voice.

RESEARCH METHODOLOGY

WEBPAGE — Text mining of project description to obtain project control variables

VIDEO →
- **VISUAL TRACK** → Detect human faces (image recognition) → Obtain control variables for visual characteristics
- **AUDIO TRACK** → Identify no. of voices (speech recognition) → Obtain control variables for linguistic and audio characteristics using text mining and audio analysis
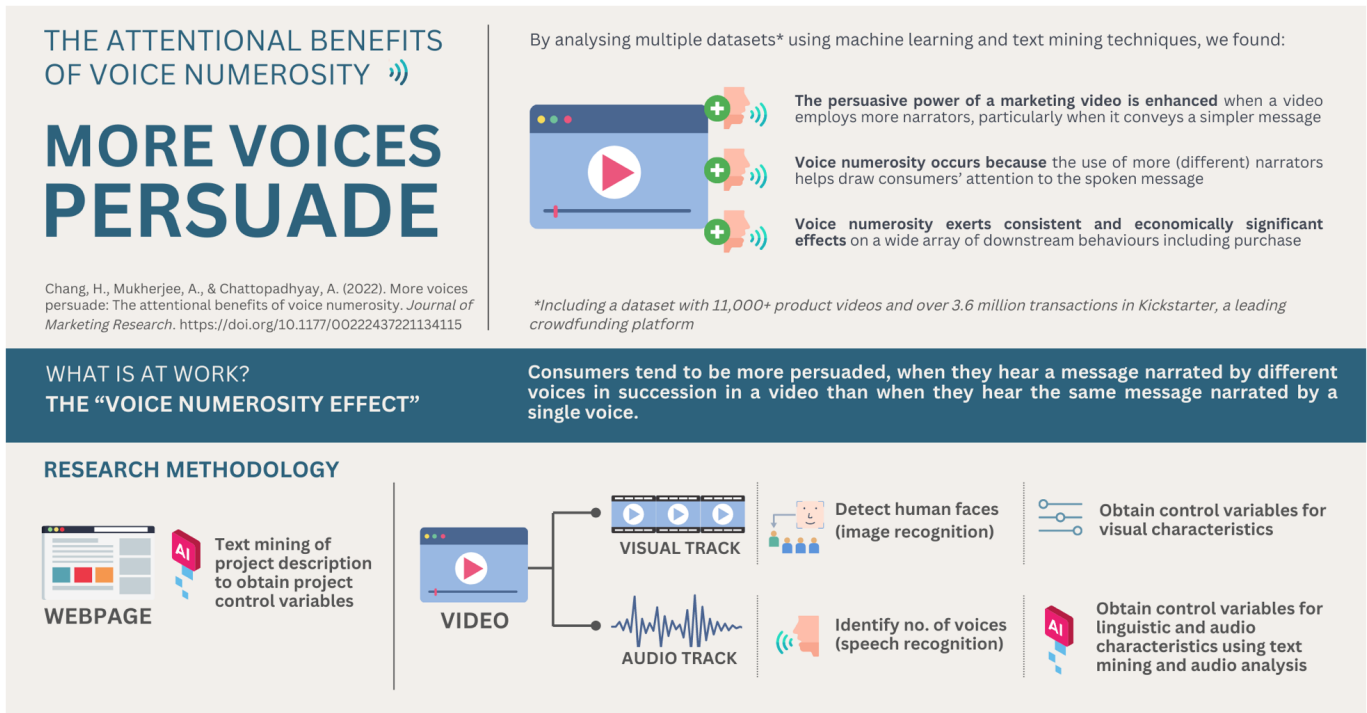
Image credits: Singapore Management University Library. Used with permission.

As presented in Figure 1, across a broad spectrum of models and datasets, we find:

- The persuasive power of a marketing video is significantly enhanced when a video employs more narrators, particularly when it conveys a simpler message. We term this phenomenon "voice numerosity".

- Voice numerosity likely occurs because the use of more (different) narrators helps draw consumers' attention to the spoken message.

- Voice numerosity exerts consistent and economically significant effects on a wide array of downstream behaviours including the likelihood of purchase and in enhancing the persuasiveness of marketing videos such as video advertisements.

- The measured effects are economically consequential. For example, having an additional voice in a project video on Kickstarter (a crowdfunding website; Dhanani and Mukherjee 2019; Mukherjee et al. 2019), ceteris paribus, is associated with (1) an

increase of about $12,795 in pledged amount (a 39% increase), (2) 118 more customers backing the project (a 38% increase), and (3) a 1.6% greater probability that the project is successfully funded (a 6.5% increase).

These findings underscore the importance of the human voice in videos as a potential strategic design element. Conversations with practitioners show that even as voice-over narration is a typical and common video presentation format, practitioners today remain unaware that due to the central role of voice in human cognition, the mere addition/subtraction of voice actors can have a significant effect on consumer attention and processing of the product message, and therefore on downstream consumer behaviour. In particular, even as current industry practice focuses on a need for "a clear speaking voice" (YouTube Advertising 2019) that signals "authority" and "relatability" (Voices 2018) it has yet to consider using number of voices as a strategic design element. Thus, as video marketing continues to affect consumers' purchase journeys, their research offers specific recommendations on voice-over narration for marketing practitioners and architects of the consumer information environment. Specifically, the findings suggest that for more difficult-to-comprehend product messages (e.g., said at three words per second), it might be more effective to have just one narrator. In contrast, for messages that are simple to comprehend (e.g., said at one word per second), it may be worthwhile to have multiple narrators, to leverage the voice numerosity effect.

Moreover, a crucial type of video advertisement in which voice-over narration is commonly employed is political advertising. The voice numerosity effect suggest that the use of many voices in a political advertisement, rather than only the voice of the candidate or any other voice narrator, is likely to affect potential voters. In addition, the attention benefits of voice numerosity are likely to most benefit political videos with simpler messages that are easier to comprehend, and messages that are geared towards constituents who are more pressed for time or otherwise have limited ability to process the message. Importantly, these factors have been shown in prior research to correlate with demographic indicators such as age, education, and location, which are identifiable and targetable by political parties and organizations.

The research also underscores the value of combining different kinds of data (audial, textual, and visual) analysis methods in a single study to tease apart the implications of many different data features on consumers (Chang and Mukherjee 2022). A video comprises of many forms of information such as the audial signature of the track (e.g., the volume of the track), the visual signature of the track (e.g., the resolution of an online ad), the key features of the spoken message (e.g., the positivity of the spoken message), and the number of off-screen vs. on-screen voices. Identifying and coding all of these factors in many videos would have been cost prohibitive. Therefore, the research employed a wide array of machine learning methods that involve the separation of a video into its audial and visual tracks, the isolation of different voices through the use of speaker diarization (a deep learning technique; Makino et al. 2019), transcribing the audial track to account for differences in audial content, and using image recognition to account for differences in visual content, amongst other methods, to derive key control variables – variables that account for all differences between two observations that may otherwise confound measurement of a focal hypothesis – which were then included in the econometric model.

For firms swamped with diverse forms of data fragmented in data silos, this research thus provides a primer on the application of computational scale methods to derive conceptually relevant variables and to apply them to the study of theoretically rich and substantively crucial research questions in ecological data – data that tracks and describes the real-world behaviour of their consumers. The approach – the use of machine learning to derive many variables at scale from unstructured data to examine key focal hypotheses – is new to the psychology and marketing literatures and presents a novel avenue for data exploration and analysis for the social sciences (Athey and Imbens 2019; Mukherjee and Chang 2022). This methodology derives from recent advances in neural networks and machine learning, and benefit from the increased availability of computational resources such as cloud computing, which are now generally available through web interfaces such as Google colab without requiring complex, specialized, dedicated, and expensive analytical resources.

Bibliography

1. Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. Journal of Personality and Social Psychology, 37(5), 715.

2. Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. Trends in cognitive sciences, 8(3), 129-135.

3. Chaiken, Shelly and Alice Eagly (1983), "Communication Modality as a Determinant of Persuasion: The Role of Communicator Salience," *Journal of Personality and Social Psychology*, 45 (2), 241–56.

4. Chang, H. H., Mukherjee, A., & Chattopadhyay, A. (2022). More Voices Persuade: The Attentional Benefits of Voice Numerosity. *Journal of Marketing Research*, forthcoming. **https://doi.org/10.1177/00222437221134115**

5. Chang, H. H., & Mukherjee, A. (2022). Using machine learning methods to extract behavioral insights from consumer data. In J. Wang (Ed.), *Encyclopaedia of Data Science and Machine Learning* (forthcoming). IGI Global.

6. Chattopadhyay, A., Dahl, D. W., Ritchie, R. J., & Shahin, K. N. (2003). Hearing voices: The impact of announcer speech characteristics on consumer response to broadcast advertising. *Journal of Consumer Psychology*, *13*(3), 198-204.

7. Dhanani, Q., & Mukherjee, A. (2019) Is Crowdfunding the Silver Bullet to Expanding Innovation in the Developing World? (accessed November 25, 2022). **https://blogs.worldbank.org/psd/crowdfunding-silver-bullet-expanding-innovation-developing-world**.

8. Grewal, R., Gupta, S., & Hamilton, R. (2021). Marketing insights from multimedia data: text, image, audio, and video. *Journal of Marketing Research*, 58(6), 1025-1033.

9. Horowitz, S. S. (2012). The universal sense: How hearing shapes the mind. Bloomsbury Publishing USA.

10. Makino, T., Liao, H., Assael, Y., Shillingford, B., Garcia, B., Braga, O., & Siohan, O. (2019, December). Recurrent neural network transducer for audio-visual speech recognition. In *2019 IEEE automatic speech recognition and understanding workshop (ASRU)* (pp. 905-912). IEEE.

11. Mukherjee, A., & Chang, H. H. (2022). Describing rosé: An embedding-based method for measuring preferences. In A. Humphreys, G. Packard, & K. Gielens (Eds.), AMA winter academic conference proceedings (pp. 150-151). American Marketing Association.

12. Mukherjee, A., Chang, H. H., & Chattopadhyay, A. (2019). Crowdfunding: sharing the entrepreneurial journey. In R. W. Belk, G. M. Eckhardt, & F. Bardhi (Eds.), *Handbook of the sharing economy* (pp. 152-162). Edward Elgar Publishing.

13. Rutten, S., Santoro, R., Hervais-Adelman, A., Formisano, E., & Golestani, N. (2019). Cortical encoding of speech enhances task-relevant acoustic information. *Nature Human Behaviour*, *3*(9), 974-987.

14. Voices (2018), "2018 Voice Over Trends In Marketing and Advertising," (accessed December 9, 2022), **https://static.voices.com/assets/uploads/client/2018-Voice-Over-Trends-In-Marketing-v2.pdf**.

15. YouTube Advertising (2019), "How to Make a Video Ad for Your Marketing Strategy," (accessed December 9, 2022), **https://www.youtube.com/ads/resources/how-to-make-a-video-ad-that-fits-your-marketing-strategy/**.

Hannah H. Chang

Hannah H. Chang is Associate Professor of Marketing and Director of PhD Programmes at Singapore Management University. She examines the intersection of digital technologies, product innovations, and consumer behavior, and their strategic marketing and public policy implications. Her research has appeared in Journal of Marketing Research, Psychological Science, among others.

Anirban Mukherjee

Anirban Mukherjee is Visiting Scholar at Cornell University. He is an expert in quantitative and computational marketing. His research has appeared in leading journals, including the Journal of Marketing, Journal of Marketing Research, International Journal of Research in Marketing, Journal of Retailing, and Management Science.

Amitava Chattopadhyay

Amitava Chattopadhyay is a Professor of Marketing, the GlaxoSmithKline Chaired Professor of Corporate Innovation, and Marketing Area Chair at INSEAD. His research has appeared in leading journals, including the Journal of Marketing Research, Journal of Consumer Research, Journal of Consumer Psychology, Journal of Marketing, Marketing Science, and Management Science.