

# Would People Mumble Rap to Alexa?

Casey C. Bennett\*

Department of Data Science  
Hanyang University  
Seoul, Korea  
cabennet@hanyang.ac.kr

Minha Lee

Department of Industrial Design  
Eindhoven University of Technology  
The Netherlands  
M.Lee@tue.nl

## ABSTRACT

Unclear speech, like mumbling, is difficult to understand for people, and even harder for conversational user interfaces (CUI) to process. Yet, there are multiple reasons why unintelligible speech is meaningful between humans, playing a critical role in social dynamics and status signaling that evolved in humans over time allowing us to form cohesive social groups for survival. For example in modern times, humans often use such *changed speech* in order to make themselves understandable only by in-group members, e.g. “mumble rap”, while subtly excluding out-group members. As such, we argue here that future CUIs must be attentive to *how* people use various forms of non-standard changed speech (e.g. mumbling, dialect, slang, inflection) to express themselves, lest CUIs be socially inept. Based on psychological, linguistic, and cross-cultural research, we point out several major challenges for researchers: 1) current CUIs typically omit non-standard speech like mumbling which are critical to human social communication, and 2) in the future humans may innately form ingroups with their personal CUIs resulting in speech behaviors meant to exclude outgroup members (both humans and machines). Both of those challenges require more research to address. Moreover, the use of changed speech for status signaling and ingroup/outgroup (IG/OG) signaling appears to be a phenomenon that varies across diverse cultures, languages, and situations, which CUI designers and engineers need to be mindful of going forward.

## CCS CONCEPTS

- **Human-centered computing** → **Natural language interfaces**;
- **Computing methodologies** → **Natural language processing**;
- **Computer systems organization** → *Robotics*.

## KEYWORDS

Language, Ingroup/Outgroup, Signaling, Speech inflection, Artificial agents

### ACM Reference Format:

Casey C. Bennett and Minha Lee. 2023. Would People Mumble Rap to Alexa?. In *ACM conference on Conversational User Interfaces (CUI '23)*, July 19–21, 2023, Eindhoven, Netherlands. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3571884.3603757>

\*Corresponding Author

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*CUI '23*, July 19–21, 2023, Eindhoven, Netherlands

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0014-9/23/07.

<https://doi.org/10.1145/3571884.3603757>

## 1 INTRODUCTION

### 1.1 Overview

We have all had the experience of going to a job interview, or other formal settings, where we had to clearly articulate our speech in order to present ourselves as polite or educated, in order to make sure we were understood. That contrasts with how we might speak with our friends in more casual settings, where we may adopt different patterns of speech inflection, intonation, rhythm, slang, and so forth, sometimes to the point of being barely understandable to others [14, 19]. But that raises an intriguing question: **why would humans ever communicate but not do so clearly?** If we were not going to be understood, would it not be easier to just not talk at all?

That question suggests that there perhaps may be an advantage to unclear communication in certain settings, something that goes beyond the speech content itself, which has implications for the development of artificial speech agents and conversational user interfaces (CUI) in the future. For example, in human-human speech, people often mumble or even omit words because the content of the message is not always as important as the way it is conveyed [16, 17]. The ability of the listener to infer the speaker’s intent when information is missing (e.g. entire words) is one of the aspects that distinguishes fluent speakers from non-fluent second language learners [21, 34]. In this position paper, we make the argument that focusing purely on 100% comprehension with CUI misses the other “signaling” aspects of human-human communication that allow us to form cohesive groups to cooperate on tasks. In particular, we focus on the phenomenon of “mumbling” in human speech, with an example of mumble rap, and the lessons it can provide for us toward the development of more immersive user experiences with conversational agents.

### 1.2 What is mumbling?

Mumbling in general refers to a lack of clear articulation during speech, and can be thought of as part of a broader phenomenon of various kinds of speech inflection and changes in rhythm/intonation that speakers adopt in certain situations. One useful example is a genre of music known as “mumble rap”, which illustrates some of socio-cultural aspects of mumbling.

Mumble rap, as a sub-genre of rap, is most often characterized by inarticulate lyricism nested within polyphonic soundscapes, but also refers to a generation of hip hop artists that often self-publicized their work online first on platforms like Soundcloud with “raw” recording quality <sup>1</sup> [5, 6, 32]. Names include Lil Yachty or Migos in the U.S. to Yung Hurn in Austria; these artists’ challenge is less focused on semantic content and more on the affective style of

<sup>1</sup><http://www.soundcloud.com/>

delivery; in parallel, their global audience engages with their music as a highly integrated experience across diverse online and social media platforms to keep up with artists, such as moving across Instagram, Soundcloud, and more [13, 23, 36]. Hence against the established backdrop of hip hop as a form of political expression of the oppressed claiming their linguistic identity and empowerment, mumble rap is interpreted to have emerged as a further act of subversion, prioritizing performative delivery (with online-offline fluidity) rather than lyrical mastery, going against the “clean”, articulate use of spoken language, resisting white supremacy [13, 32, 38]. Mumble rap internationally has a comparatively apolitical, mundane dimension on what artists choose to rap about if only lyrical content is considered [23].

There are two ways of thinking about group identification in mumble rap, i.e., 1) the broader oppressed vs. oppressor dichotomy that was already present in the history of hip hop, but newly 2) the post- vs. pre- Soundcloud generation of rappers, in which “mumble rap” was first used as a derogatory term by the “old guard” to refer to the younger generation of artists that are comfortable mixing autotune and phonetic reduction in speech (such as “should have” to shoulda). The focus is on their creative act of delivery that surpasses the online-offline divide as “digital natives” establish themselves on social media and music platforms to produce and consume music. Thus, what is communicated is beyond the literal semantic content, i.e., *what*, to the meaning-making at the performance level of rap, which forms a generational movement through and with technology, be it in the production, circulation, or consumption (such as autotune and Soundcloud) of rap i.e., *how*. The establishment they fight is no longer just white supremacy<sup>2</sup>, but established rappers that dictate what consists of the “right” way to rap.

### 1.3 Language and InGroup/OutGroup Signaling

The previous section illustrates that speech alteration, such as mumbling, can often be used by humans as a form of signaling group identity or status, distinguishing “us vs. them”. That can be more broadly connected to the concept of ingroup vs outgroup (IG/OG) psychology and its well-known effects on human behavior. In short, humans like many primates are highly social animals that depend on cooperation for survival. IG/OG psychology is thus thought to be core component of how we evolved to form cohesive groups [3, 18, 27]. For ancient humans during most of our evolutionary history, being kicked out of the tribe meant death, and strangers from outside the tribe represented potential danger.

Given that, it became necessary for humans to use *status signaling* to indicate both group membership and our relative position within the group. As detailed in Section 2 below, that has been the focus of much research in recent years. Some of that research also focuses on how language in particular is connected to IG/OG related signaling, including use of “covert signaling” through subtle changes in speech [35].



**Figure 1: Future “mumble rapper” CUI robot, as generated by Bing (Dall-E) Commons. (<https://www.bing.com/images/create>).**

## 2 RESEARCH EVIDENCE

### 2.1 Psychological Research

As mentioned above, there is significant research into IG/OG signaling within the fields of psychology and human evolution, related to how humans form cohesive social groups to survive [3, 18, 27]. Those groups typically use markers, either visual or linguistic, to distinguish themselves from “others”. Indeed, even the largest social groupings amongst humans (e.g. nations, tribes) are often demarcated by language or dialect boundary [31]. Interestingly, there appear to be antecedents for such “us vs. them” behavior in other primates, such as macaques and chimpanzees [25]. For example, chimpanzee groups are known to use unique “pant hoots” to distinguish groups from one another, particularly when different groups are living close together [9].

Similarly in humans, such ingroup biases are present in small children at adult-like levels even early on in their development before any formal schooling, indicating that IG/OG is an ancient innate component of human survival strategy embedded within our behavior via evolution [12]. Such ingroup biases appear to have a neurobiological basis, involving modulation of oxytocin release that increases trust and empathy towards the ingroup (or signals related to it) while dampening it toward the outgroup [11].

The above psychological aspects manifest in human behavior often in the form of *signaling*. Such signaling may take the form of visual status symbols (e.g. a crown or fancy watch) or verbal communication [8]. For the latter, such verbal signaling may involve the use of different vocabulary, grammar structures, or manners of speech (e.g. accents). In some situations, humans even use subtle

<sup>2</sup><https://www.theatlantic.com/politics/archive/2017/10/the-language-of-white-supremacy/542148/>

forms of *covert signaling* that are deliberately intended to be detectable only by the in-group, but oblivious to the out-group [35]. Any of us who have raised teenagers certainly have witnessed their use of esoteric slang to communicate amongst themselves without nearby adults fully understanding (e.g. "No Cap", "YEET"). In general, covert signaling is particularly interesting in that it allows humans to signal status and identity to other in-group members in subtle ways that may be beneficial when individuals are members of multiple overlapping but distinct groups. That maintains the ability for individuals to still cooperate with other groups later but not alienate their primary in-group, which may be advantageous [35]. We know that signaling is not just about group membership identity, but also maintenance of one's status within their in-group, even amongst chimpanzees [27]. Mumble rap (and mumbling in general) may in fact be one form of covert signaling among humans.

## 2.2 Linguistic Research

There is ample research from linguistics that may shed some light on how exactly mumbling is a form of covert signaling. In particular, studies of L2 second language learners have shown they tend to *over-compensate* their words when speaking in their second non-native language and exhibit more rigidity in pronunciation until they reach a certain level of fluency (even though their pronunciation may not be entirely correct). This is in contrast to native speakers, who exhibit more fluid pronunciation that can be seen as *under-compensation*, which impacts both their speaking and listening skills [24]. Interestingly, L2 speakers also often have difficulties when interacting with CUI agents for the same reason [40], as do children [28].

For our purposes here though, such under-compensation and over-compensation may help us understand the role of mumbling as a form of covert signaling. Given the above, it makes sense that native speakers of a language or dialect may in some situations purposely (or subconsciously) "under-compensate" their speech in order to exclude nearby non-native out-group members who speak a similar language/dialect [15]. In other words, they can subtly exclude the out-group members without seeming like they are, by making their speech more difficult to understand for anyone below native-level fluency.

Fluent-level bilingual speakers are an interesting argument for the above hypothesis. Existing research has shown that such bilingual speakers often report feeling "schizophrenic" when code switching between two of their languages, as if they were an entirely different person [33]. That suggests language, and the speech production thereof, is deeply embedded in humans' cognitive processes [2]. Beyond that, we know some forms of speech may be associated with kinds of self-identity, e.g. masculinity, and thus serve as important markers within our own sense of self [7, 29]. Indeed, research has actually shown mumbling is more common amongst men, and thus there could be argument made that mumbling is somehow associated with perceptions of masculinity [39]. In a similar vein, there is other research showing that mumbling is a form of stress-response in some individuals due to various forms of distress, e.g. PTSD [10]. Suffice it to say, the act of mumbling seems to be a behavioral response to a person's lived experienced, manifested through changes in speech production.

## 2.3 Cross-Cultural Research

Of course, one must also take into account that there are cultural differences in the way different groups of humans signal status and identity. The clearest example of this is the difference between East Asian languages (e.g. Korean, Japanese) and Western languages (e.g. English, German). In short, East Asian languages are high-context languages, requiring the listener to use more context to disentangle otherwise ambiguous communication [26]. In such languages, status and group identity are directly embedded into the language grammatically, which can alter the meaning of a spoken sentence in various ways. That is in contrast to Western languages like English that are low-context, requiring the listener to do less guessing based on context [41]. From a more personal standpoint, the authors of this paper know those differences well, as we are all fluent speakers of both English and Korean.

Interestingly, research in Western cultures across European languages has shown that individuals still use vocal cues to determine the status of the speaker (e.g. white collar vs. blue collar) [4], so it would be incorrect to infer that status is unimportant cross-culturally. Rather, a better interpretation may be that how status signaling manifests *differs* across cultures. Moreover, other research has shown that sometimes the exact same verbal signal in one culture may be interpreted differently in another culture [37]. This indicates that there is a significant cultural component to understanding how humans signal status and group identity. However, whether that affects the use of specific speech changes such as mumbling by different groups of humans is something that requires further study. There is currently a dearth of research on the subject.

## 3 DISCUSSION

While CUIs like Alexa are now designed for understanding clearly written or spoken speech, understanding intentionally unclear speech that is full of contextual cues and meaning is an open challenge. We used the example of mumbling and mumble rap for communication in order to illustrate how language and culture evolve in often unexpected ways. Subsequently, how future CUIs should then be designed as artifacts that humans interact with and grow with needs to be mindful of changes in expressive speech that "break the rules" of typical spoken language. Which leads us to our first provocative point:

- **Provocation #1:** The current approach to the design of CUI systems often omits non-standard forms of language (e.g. mumbling), which are critical parts of human social communication.

After all, the whole point of language is *communication*, not the rules or grammar used to accomplish it. This brings us back to our main question from Section 1 ... why do people mumble in the first place? Why would humans ever communicate but not do so clearly? The research evidence discussed in this paper suggest possible answers to those apparent contradictions. Perhaps people would mumble-rap to their Alexa, or other CUIs, in certain settings and scenarios (e.g. for IG/OG signaling). Understanding the reasons why they might do so has deep implications for how we think about the design of CUI in the future, which was brings us to our second provocative point:

- **Provocation #2:** People in the future will form an "in-group" with their CUI devices, like Alexa, and engage in speech behaviors meant to exclude other people.

If that indeed happens, it may necessitate CUI systems to behave differently than currently envisioned, with a focus on more subtle cues rather than the speech content itself, in order for the CUI to respond in socially-appropriate "in-group" ways. It also suggests an urgent need for more research on how *groups* of people interact with CUIs, rather than the common trend of one-on-one speech interaction research. As pointed out in Section 2, IG/OG psychology is deeply embedded in human behavior, and thus CUIs may be best viewed as a sort of "extension" of the social group. That may lead to humans using particular speech behaviors (e.g. mumbling, dialect use, intonation) with different CUIs depending on whether the CUI is viewed as part of the in-group or out-group, or even to purposely exclude nearby out-group humans from the conversation. Though the prevalence of such phenomenon may vary from one location to the next, which brings us to our third provocation point:

- **Provocation #3:** The use of status signaling cues will vary across culture and situation, and without understanding such variation CUI systems will be socially-inept.

All human speech, after all, is inherently *social* behavior. And if we assume the framework of Clifford Nass, all CUIs are thus "social actors" [30]. Lest we intend to build socially-inept CUI systems to deploy into the world, CUI must be aware of the *socially-situated context* in which they inhabit. This connects to the broader fields of human-computer interaction (HCI) and human-robot interaction (HRI), which has spent the past couple decades revealing how variation across culture, gender, and situation (e.g. work environment vs. home setting) impact human interactions with technology [1, 20, 22]. In short, what makes sense in one environment may make *zero sense* in another environment. That affects everything from robots to mobile devices to digital user interfaces, and is certain to impact CUIs as well. To our point here, it is very likely that changed speech behaviors (e.g. mumbling, dialect use) related to status and/or identity signaling will be used differently at different times by different people and by different cultures (see Section 2.3). Thus, for CUIs to be truly "socially intelligent", that will require them to understand the contexts within which those speech behaviors are appropriate, both in terms of understanding the social cues the human is attempting to convey to the CUI as well as potentially even the CUI utilizing changed speech itself.

There are a number of potential solutions to the challenges described in the above provocation points. First, the CUI community could pursue more research with CUIs *deliberately* designed to utilize changed speech (e.g. mumbling, dialect), the extent of which could be modulated (or even completely turned on/off) for experimental reasons. That may include both the production of changed speech by the CUI, as well as giving the CUI the ability to understand human uses of changed speech. An additional question may be whether it is necessary or not to utilize localized machine learning models (versus universal ones) to adapt CUIs to engage with specific in-groups (e.g. social cliques at work or school). Second, the community could pursue more cross-cultural studies of *parallel* CUI speech systems in multiple languages, i.e. systems designed to produce roughly the same semantic "meaning" in similar scenarios

across two or more languages despite grammatical/idiomatic differences in semantic construction that would cause direct translations to fail [2]. Such an approach would allow for direct comparisons of verbal status signaling across cultures (including mumbling) in the context of CUI interaction, which could later be used to create culturally-specific forms of changed speech onboard CUIs. Similarly, another intriguing avenue could be to explore CUI interactions with human bilingual speakers, to investigate whether language "code-switching" initiates shifts in signaling behavior. Third, the community could pursue experiments designed to investigate human-CUI interaction scenarios where emergent IG/OG behavior may be triggered. There is already a large body of research on human-human interaction scenarios related to such behavior from the psychology field, which could be used as templates for CUI research. In particular, we need more research involving multiple humans interacting at the same time with a CUI, or potentially multiple CUIs, to investigate group behavior dynamics in those situations.

There may of course be other possible solutions as well, so the above represent just a few starting points for further discussion. Regardless, any potential solution, including those above, will likely require multiple research steps to address ... what precisely those steps may entail is a debate we leave for the CUI community to have.

## ACKNOWLEDGMENTS

This work draws from research across many different projects, including a grant from the National Research Foundation of Korea (grant# 2021R1G1A1003801). We would also like to thank our various research collaborators from over the years.

## REFERENCES

- [1] Casey C. Bennett, Young-Ho Bae, Jun Hyung Yoon, Yejin Chae, Eunseo Yoon, Seun Lee, Uijae Ryu, Say Young Kim, and Benjamin Weiss. 2023. Effects of Cross-Cultural Language Differences on Social Cognition During Human-Agent Interaction in Cooperative Game Environments. *Computer Speech & Language* 81 (2023), 101521.
- [2] Casey C. Bennett, Say Young Kim, and Benjamin Weiss. 2023. Bilingual Robots to Explore Cognitive Effects of Bilingualism thru Human-Robot Interaction. *CUI@HRI Workshop at the ACM/IEEE Conference on Human-Robot Interaction (HRI'23)* (2023), 1–4.
- [3] Jonah Berger and Chip Heath. 2008. Who drives divergence? Identity signaling, outgroup dissimilarity, and the abandonment of cultural tastes. *Journal of personality and social psychology* 95, 3 (2008), 593.
- [4] Bruce L. Brown and Wallace E. Lambert. 1976. A cross-cultural study of social status markers in speech. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* 8, 1 (1976), 39.
- [5] Preezy Brown. 2023. Lil Yachty Wants To Be Seen As More Than A "Mumble Rapper" — vibe.com. <https://www.vibe.com/music/music-news/lil-yachty-mumble-rapper-1234731145/>. [Accessed 08-Apr-2023].
- [6] Justin Charity. 2017. Declaring a Moratorium on the Term "Mumble Rap" — theringer.com. <https://www.theringer.com/2017/4/18/16047136/mumble-rap-moratorium-lil-yachty-designer-15bc3647dae9>. [Accessed 08-Apr-2023].
- [7] Antonio Chella, Arianna Pipitone, Alain Morin, and Famira Racy. 2020. Developing self-awareness in robots via inner speech. *Frontiers in Robotics and AI* 7 (2020), 16.
- [8] Emma Cohen. 2012. The evolution of tag-based cooperation in humans: The case for accent. *Current Anthropology* 53, 5 (2012), 588–616.
- [9] Catherine Crockford, Ilka Herbinger, Linda Vigilant, and Christophe Boesch. 2004. Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology* 110, 3 (2004), 221–243.
- [10] Timothy K. Daugherty and Alexander J. Mannebach. 2023. Mumbling as a Potential Marker of Posttraumatic Distress. *International Journal of Psychology and Educational Studies* 10, 1 (2023), 189–195.

- [11] Carsten K.W. De Dreu, Lindred L. Greer, Gerben A. Van Kleef, Shaul Shalvi, and Michel J.J. Handgraaf. 2011. Oxytocin promotes human ethnocentrism. *Proceedings of the National Academy of Sciences* 108, 4 (2011), 1262–1266.
- [12] Yarrow Dunham, Andrew S. Baron, and Mahzarin R. Banaji. 2008. The development of implicit intergroup cognition. *Trends in cognitive sciences* 12, 7 (2008), 248–253.
- [13] Abraham Elsher. 2018. The problem with mumble rap: Stigmatization of variant production in contemporary mainstream hip-hop. *Undergraduate thesis, Swarthmore College* (2018). <http://hdl.handle.net/10066/2000>
- [14] Mirjam Ernestus, Iris Hanique, and Erik Verboom. 2015. The effect of speech situation on the occurrence of reduced word pronunciation variants. *Journal of Phonetics* 48 (2015), 60–75.
- [15] Magdalena Formanowicz and Caterina Saitner. 2020. Sounding strange (r): Origins, consequences, and boundary conditions of sociophonetic discrimination. *Journal of Language and Social Psychology* 39, 1 (2020), 4–21.
- [16] Kenneth S. Goodman and Frederick V. Gollasch. 1980. Word omissions: Deliberate and non-deliberate. *Reading Research Quarterly* (1980), 6–31.
- [17] Doris A. Graber. 1989. Content and meaning: What's it all about. *American Behavioral Scientist* 33, 2 (1989), 144–152.
- [18] Norbert L. Kerr and John M. Levine. 2008. The detection of social exclusion: Evolution and beyond. *Group Dynamics: Theory, Research, and Practice* 12, 1 (2008), 39.
- [19] Kim Koppen, Mirjam Ernestus, and Margot van Mulken. 2019. The influence of social distance on speech behavior: Formality variation in casual speech. *Corpus Linguistics and Linguistic Theory* 15, 1 (2019), 139–165.
- [20] Leantros Kyriakoullis and Panayiotis Zaphiris. 2016. Culture and HCI: a review of recent cultural studies in HCI and social networks. *Universal Access in the Information Society* 15 (2016), 629–642.
- [21] Shiri Lev-Ari. 2015. Adjusting the manner of language processing to the social context: Attention allocation during interactions with non-native speakers. *Attention and vision in language processing* (2015), 185–195.
- [22] Velvetina Lim, Maki Rooksby, and Emily S. Cross. 2021. Social robots on a global stage: establishing a role for culture during human–robot interaction. *International Journal of Social Robotics* 13, 6 (2021), 1307–1333.
- [23] Jin Liu. 2021. Language, identity and unintelligibility: A case study of the rap group Higher Brothers. *East Asian Journal of Popular Culture* 7, 1 (2021), 43–59.
- [24] Miquel Llopart and Eva Reinisch. 2019. Robustness of phonolexical representations relates to phonetic flexibility for difficult second language sound contrasts. *Bilingualism: Language and Cognition* 22, 5 (2019), 1085–1100.
- [25] Neha Mahajan, Margaret A. Martinez, Natashya L Gutierrez, Gil Diesendruck, Mahzarin R Banaji, and Laurie R Santos. 2011. The evolution of intergroup bias: perceptions and attitudes in rhesus macaques. *Journal of personality and social psychology* 100, 3 (2011), 387.
- [26] Erin Meyer. 2014. *The culture map: Breaking through the invisible boundaries of global business*. Public Affairs.
- [27] Mark W. Moffett. 2013. Human identity and the evolution of societies. *Human Nature* 24 (2013), 219–267.
- [28] Ivonne Monarca, Franceli L. Cibrian, Angel Mendoza, Gillian Hayes, and Monica Tentori. 2020. Why doesn't the conversational agent understand me? a language analysis of children speech. In *Adjunct proceedings of the 2020 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2020 ACM international symposium on wearable computers*. 90–93.
- [29] Alain Morin. 2003. Inner speech and conscious experience. *Science and Consciousness Review* 4 (2003).
- [30] Clifford Nass, Jonathan Steuer, and Ellen R. Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 72–78.
- [31] Daniel Nettle and Robin I.M. Dunbar. 1997. Social markers and the evolution of reciprocal exchange. *Current Anthropology* 38, 1 (1997), 93–99.
- [32] Delon A. Omrow. 2018. Mumble in the Bronx: A socio-cultural analysis of 'mumble rap' as environmental discourse. *European Journal of Research and Reflection in Arts and Humanities Vol* 6, 1 (2018).
- [33] Aneta Pavlenko. 2014. *The bilingual mind: And what it tells us about language and thought*. Cambridge University Press.
- [34] Michael Rost and Steven Brown. 2022. Second language listening. In *Handbook of practical second language teaching and learning*. Routledge, 238–255.
- [35] Paul E. Smaldino, Thomas J. Flamson, and Richard McElreath. 2018. The evolution of covert signaling. *Scientific reports* 8, 1 (2018), 1–10.
- [36] Bernhard Steinbrecher and Michaela Pichler. 2021. “That’s the Kind of Music You Hear on the Subway without Headphones!”: Musical Evaluation Practices of Adolescents in Vienna. *Popular Music and Society* 44, 5 (2021), 579–609.
- [37] Jessica L. Tracy, Azim F. Shariff, Wanying Zhao, and Joseph Henrich. 2013. Cross-cultural evidence that the nonverbal expression of pride is an automatic status signal. *Journal of Experimental Psychology: General* 142, 1 (2013), 163.
- [38] Valentin Voloshinov. 1929. Marxism and the Philosophy of Language. In *Literary Theory: An Anthology*, J. Rivkin (Ed.). London: Blackwell Press, 45–108.
- [39] Melanie Weirich, Susanne Fuchs, Adrian Simpson, Ralf Winkler, and Pascal Perrier. 2016. Mumbling: Macho or morphology? *Journal of Speech, Language, and Hearing Research* 59, 6 (2016), S1587–S1595.
- [40] Yunhan Wu, Daniel Rough, Anna Bleakley, Justin Edwards, Orla Cooney, Philip R. Doyle, Leigh Clark, and Benjamin R. Cowan. 2020. See what I’m saying? Comparing intelligent personal assistant use for native and non-native language speakers. In *22nd international conference on human-computer interaction with mobile devices and services*. 1–9.
- [41] June O. Yum. 1988. The impact of Confucianism on interpersonal relationships and communication patterns in East Asia. *Communications Monographs* 55, 4 (1988), 374–388.