DOI: 10.17469/O2111AISV000015

DUCCIO PICCARDI

Reference, sound symbolism, and variation. In search of an elusive link

This paper deals with an unresolved issue concerning Ohala's frequency code proposal on the origins of sound symbolism. To date, there is a lack of evidence supporting the assumed human capability of extracting meaning features from the indexical realm and repurposing them in sound-symbolic denotation strategies. Here an experimental replication of this mechanism was attempted, built on previous observations that, in the Florentine variety, long-lag voiceless plosives are invested with indexical meanings related to manliness. Through a forced-choice task, 22 local senior high-school students were asked to match nonwords with or without initial long-lag voiceless plosives with male or female specimens of fictitious animal species. Results failed to highlight the expected association between aspirated nonwords and male specimens; however, stimuli read by a male (vs. female) voice were significantly more likely to be matched with male animals, suggesting that sound-symbolic transfers between meaning domains can indeed be triggered under specific circumstances.

Keywords: sound symbolism, denotation, indexicality, frequency code, Voice Onset Time.

1. *Sound symbolism: from maluma to hero*

About a century ago, almost contemporarily, two researchers of very different backgrounds made us aware that, indeed, people actively follow the principle that sounds can bear specific meanings and associations to referential features, which was previously only hinted in anecdotes, theoretical recounts, or through lexical analyses (Magnus, 2013). On the one hand, involving more than 500 participants, anthropologist-linguist Sapir (1929) tested the hypothesis that vowel quality evokes referential size (*mal-mil* effect); on the other, psychologist Köhler (19305: 242-243), elaborating on Usnadze (1924), showed his audience two abstract figures, one being spikey, the other curvilinear, while elusively stating that "if [...] the reader is asked to choose which he would rather call *takete* or *maluma*, he will be probably be able to decide with ease". The success of two major schools viewing the experimental tradition on what we will call from now on "sound symbolism" as outside their primary research interests slowed down its recognition in the mainstream disciplinary scenario (Magnus, 2013): Saussurian structuralism, with its centrality of arbitrariness, and generativism, with its focus on Grammar and its troubled relationship with stochastics and semantics (Kawahara, 2020; see also

Alderete, Kochetov, 2017)[1]. As a consequence, reports have begun to pop up on the quite strict (Svantesson, 2017), quasi-demonizing (Nobile, 2019) marginalization of sound symbolism in the academia.

In recent times, several studies underlined that sound symbolism is currently enjoying a second youth, which is signalled through a steady increase in the related publications (*ibid.*; Sidhu, Pexman, 2018; Nielsen, Dingemanse, 2021). Adopting a rather pragmatic mindset, Shinohara, Uno (2022) tried to interpret this novel contingency. From their viewpoint, sound symbolism succeeded in overcoming disciplinary resistances since it was proven to be a valuable asset in refining our knowledge about crucial research topics, such as the organization of the lexicon, language acquisition, and its origins. These findings, in turn, prompted to relativize "the Saussurian dictum of arbitrariness" *(ibid.*: 2). Nobile (2019) looked at this change of attitude through wider sociological lens. In his recount, Saussurian arbitrariness of the sign is seen as a postulate of demarcation between the human and the natural sciences, or, latu senso, between (rational) culture and (physical) nature. However, in contemporary times, economic and environmental crises induce scientific paradigms to rethink man as essentially bound to the rules of the physical world. This may explain the surge of interest in embodiment and enaction in the cognitive disciplines, and a greater tolerance to sound symbolism (i.e., non-arbitrariness) in linguistics. In this context, it comes as no surprise that researchers are willing to find a legitimate place for sound symbolism in their disciplinary milieu. An extreme case in point is Kawahara's (2020) attempt at promoting it among theoretical phonologists, which provides insights on why researchers trained in the generative tradition should care about this topic. Indeed, recent disciplinary trends in phonology, such as a renewed attention to phonetic events (Hannahs, Bosch, 2018), as well as to gradient and probabilistic phenomena (Alderete, Finley, 2023) seem to create new opportunities for this field to develop an interest in sound symbolism.

In this paper, I will deal with another field which is currently striving to legitimate sound symbolism in its tradition: variationist sociolinguistics (§ 1.1). In particular, by addressing the underspecified relationship between indexicality and denotational meaning in a renowned model on the origins of sound symbolism (the frequency code, § 1.2), I will join the contemporary interest in closing theoretical gaps (e.g., Sidhu, Pexman, 2018; Ekström, 2022). In order to do that, I will propose a forced-choice experiment (§ 4) to 22 Florentine Italian students under the assumption that, since long-lag voiceless plosives are locally invested with social meanings related to manliness (§ 2), they should be able to sound-symbolically extract this feature and

[1] At the risk of oversimplifying, generativism traditionally placed probabilistic tendencies outside linguistic competence, while dichotomous, categorical distinctions (grammatical vs. ungrammatical) are pertinent to grammar (see also Alderete, Finley, 2023). Moreover, the Chomskyan tradition gave a marginal role to semantics (Piattelli-Palmarini, Cecchetto, 1997), starting with the skepticisms of Chomsky himself. See Glanzberg (2021) for a historical overview on semantics in Chomskyan work, especially with respect to the autonomy of syntax, and also Krivochen, Lacková (forthcoming)'s discussion on Chomskyan rejection of externalism as a potential problem for the analysis of iconicity in the generative framework.

coherently use it in denotation. Methodological considerations are expounded in § 3, while the results are discussed in § 5. Lastly, conclusions and directions for future research are drawn in § 6.

1.1 The social life of sound symbolism

The disciplinary intersection between sound symbolism and sociolinguistics is, again, one of initial skepticism and recent enthusiasm. An early systematic attempt at finding a place for it in the Labovian paradigm can be read in Gordon, Heath (1998). The authors argue that the role of speaker sex in language variation and change can be traced back to sound-symbolic preferences based on the dimorphism of the phonatory apparatus (Ohala, 1984; more on this in § 1.2). In other words, women are expected to be leaders in raising/fronting, while men in lowering/backing phenomena. In the comments reported after the main text of the paper, this proposal was deemed as simplistic at best, and its predictions were consequently never systematically tested for (Eckert, 2019: 764). Nonetheless, Labov himself picked up this idea with interest, while ultimately dismissing it because of too many counterexamples (Labov, 2001: 291-292; but see the reply in Heath, 2003) and its determinism (Eckert, Labov, 2017). It should be noted that Gordon and Heath's (1998) theory was published in an anthropology journal. Indeed, at that time, anthropological linguistics had already entertained fruitful discussions on the roles of sound symbolism, especially in the expression of affect (e.g., Samarin, 1978; Ochs, Schieffelin, 1989; Besnier, 1990). These reflections will have a clear echo in the more recent inclusion of sound symbolism in the sociolinguistic mainstream.

Formalized in Eckert (2012), the third wave of variationist sociolinguistics has clear ethnographic origins (*ibid*.: 93). This successful approach shifted the focus of variation to individual stylistic practices, meaning-making processes, and indexical mutability. The study of socio-semiotics broadens the scope of the sociolinguistic inquiry beyond the realm of language per se. Retracing the steps of indexical elaborations, sociolinguistics ends up becoming an embodied discipline concerned with discourse as part of the material world (Bucholtz, Hall, 2016). Mirroring Nobile's (2019) considerations (§ 1), this drift coherently led to a rediscovery of sound symbolism. In Eckert (2010), the author explains the affective use of backed/fronted vowel variants enacted by two preadolescent girls as the result of contextual indexical connections stemming from size sound symbolism, which "lies somewhere between the natural and the conventional" (*ibid.*: 70). Lee (2020) expanded our dictionary of social meanings flourishing from size sound symbolism noting that vowel frontness is perceived as "refined" by speakers of Baba Malay. Similar perspectives are adopted for the explanation of consonant phenomena, such as sibilant variation (Levon, Maegaard & Pharao, 2017). Thus, sound symbolism is now included in seminal overviews as a fundamental indexical resource (e.g., Eckert, 2019; D'Onofrio, Eckert, 2021) and, as Woschitz (2019) argutely observed, its acceptance in sociolinguistic theory signals the emergence of interferences with Labovian standpoints, such as the assumed social meaninglessness of specific

variation (and change) phenomena. Moreover, the fact that individual sound-symbolic resources can point to both the qualities of the denotatum (e.g., [i] means small object) and the speaker alike (e.g., [i] means small utterer) strengthens an emerging sociolinguistic view (e.g., Johnstone, 2010) positing fluid boundaries between the referential and the social meaning domains (D'Onofrio, Eckert, 2021: 30; but see Munson, 2010).

Being framed in the larger picture of iconicity as the ("conventional but [...] invok[ing] the natural") property of a sign whose "vehicle is construed as resembling its object in some diagrammatic or sensual way" (D'Onofrio, Eckert, 2021: 33), sociolinguistic sound symbolism found its ideal phylogenetic framework in Ohala's (1984) "ethological" theory, the frequency code, which is grounded in animal and human physiology. In the next paragraph, I will focus on the frequency code, interpreting it as a theory of indexicality which crucially underspecifies its presumed relationship with denotational meaning.

1.2 Disentangling the frequency code *patola*

In its original formulation, Ohala's (1984) frequency code theory is presented as a Gujarati *patola* fabric, and each of its arguments as a colorful thread which should be evaluated only after being woven into the pre-planned pattern. This rhetorical structure makes the frequency code a "highly synergetic", "synthetic", "bold" proposal (Winter, Oh, Hübscher, Idemaru, Brown, Prieto & Grawunder, 2021: 2) in need of further inquiries.

Ohala posited that physical correlations found in the animal world, e.g., the direct correspondence between the overall size of a living being and the mass of its vibrating membranes ("vocal cords in mammals, syrinx in birds"; Ohala, 1984: 5), and the inverse correspondence between the latter and the height of the frequencies they can produce, originated several linguistic phenomena, including social indexing and size (referential) sound symbolism. By disposing Ohala's threads in a causal order, we can infer that frequency as a feature carrying biologically determined indexicality (i.e., signaling body size; Foulkes, 2010) was involved in basic (socio) semiotic reinterpretations, and thus contextually used to enact hostile stances and exploit the concept of size in interindividual competition. These prompted a sexual selection in the human species, as successful males manifested enlarged phonatory apparatuses capable of enhancing their aggressive displays, and ended up with the establishment of our sexual dimorphism, which in turn added a new layer of sex-based biologically determined indexicality to human voice (i.e., signaling speaker sex through different average F_0). Thus, Ohala's theory concerns the role of indexicality in shaping our reality, retracing several semiotic moves between the natural and the conventional. The other proposed threads concerning these facets of meaning, such as the affective uses of F_0, seemingly match the colors of the general pattern, as they can be interpreted (just as the other sociolinguistic phenomena expounded above, § 1.1) as the continuation of an indexical order (Eckert, 2008 and references therein) stemming from general co-occurrences regarding frequency. On the contrary, the

addition of referential meaning (i.e., tonal and segmental sound-symbolism) to the pattern stands out as a discolored motif. The formation of sound-symbolic mappings onto referential features lies somewhere in the course of this indexical process, but what triggered this osmosis between meaning domains is not entirely clear. Ohala (1984: 5-6) is aware that the explanation of this type of sound symbolism is "more problematic" in his theory, since speakers using frequency resources for the denotation of size are "presumably not trying to appear small or large [...], or even submissive or dominant". The author only briefly addresses this point by stating that frequency serves as a multipurpose trigger of hearer reaction to a feature of the communicative environment, either being part of the utterer or the denotatum.

The vagueness of the frequency code in this regard is baffling if we consider that on the line dividing indexical from referential sound symbolism lie important disciplinary boundaries: in fact, psycholinguists explicitly question the inclusion of the former in their analytical agenda, while still mentioning Ohala's proposal among the main theories covering sound-symbolic associations (e.g., Sidhu, Pexman, 2018: 1626, 1631; Akita, 2021: 3). On the other hand, third-wave sociolinguistics (e.g., Eckert, 2010; Levon, Maegaard & Pharao, 2017) usually relies on the observations advanced in Silverstein (1994) in order to substantiate potential sound-symbolic mergers between meaning domains. Silverstein (1994) reports that Wasco-Wishram has a highly structured variation system based on sound symbolism to express diminutives/augmentatives. What is interesting here is that this system blends into the affective/socio-indexical realm, as the variants convey properties of both the referent and the speaker ("affectively engaging smallness" vs. "affectively engaging bigness"). However, note that diminutives are known to simultaneously express meanings belonging to different domains independently of their means of expression (e.g., Jurafsky, 1996; Dressler, Merlini Barbaresi, 1992 for Italian), so that the sound-symbolic vehicle of Silverstein's examples might be incidental.

Overall, finding ulterior evidence in favor of Ohala's hypothesized evolutionary process may be beneficial for both the utilization of the frequency code in psycholinguistics and the promotion of the transversal acceptation of meaning endorsed by sociolinguistics. Thus, the time seems ripe to move to an experimental approach and test if speakers can transfer meaning from the indexical to the denotational domain. In the next paragraph, I will present the social-meaning-rich feature that will be deployed in this work: Voice Onset Time.

2. *Voice Onset Time circling around hyperarticulation*

Formalized in Lisker, Abramson (1964), Voice Onset Time (VOT) is a sub-segmental feature corresponding to the lapse between the release of a voiceless plosive and the onset of periodicity of the following vowel (Francis, Ciocca & Yu, 2003). Scobbie (2006) strongly promoted the idea that VOT, in addition to its role as a phonological cue, can be invested with indexical values. Indeed, VOT is phonetically enhanced through hyperarticulation (e.g., Schertz, 2013); thus,

long-lag VOT segments represent the fortis endpoint of variation systems based on gradient articulatory strength, and the construction of social meaning on these variants has to take into account this generality (Acton, 2021: 121).

Developing the flagship representational tool of third-wave sociolinguistics, the indexical field, Eckert (2008) gathered together several reports on the social meaning attributed to hyperarticulated /t/ release in American English and disposed them in a descriptive graphical arrangement in which proximity between tags and distance from the center loosely represent plausible socio-semiotic linkages and reinterpretative moves, or increases in indexical order, without predicting fixed directions. Expanding and critically analyzing Eckert's field, Drummond, Schleef (2016: 53) noted that, in order to hold together the disparate meanings of /t/ release, it is necessary to presuppose a central indexical value shared by all the instantiations of the phenomenon, i.e., something related to hyperaticulation ("extremely articulate or clear speech").

The indexical field of hyperarticulated voiceless plosive releases (e.g., Piccardi, 2021: 58) does justice to its primacy in the development of this line of research, showing the tool's potential of providing ideological connections between otherwise irreconcilable meanings (Gafter, 2021). In fact, hyperarticulation originates the impression of both accuracy and emphasis, which can lead to antithetic high-order meanings. For example, while Karthikeyan, Puts, Aung, Link, Rosenfield, Mackiel, Casey, Marks, Cristo, Patel, Santos & Geher (2023) recently noted that, in American English, men producing aspirated /t/s in word-final position (vs. unreleased /t/s) are perceived as more prestigious than physically dominant by other men, Zimman (2023) discussed the attribution of anger to Black speakers' released /t/s.

In Piccardi (2017a), I commented on a VOT pattern found in the speech of 24 Italian Florentine speakers. In this variety, a widespread, highly salient lenition targets post-vocalic plosives (the *gorgia toscana*; e.g., Marotta, 2008); in addition to that, very few studies (e.g., Castellani, 1960) report a specular phenomenon, the *gorgia enfatica*, strengthening the voiceless plosives in initial, post-consonant and gemination contexts and turning their phonologically congruous short-lag releases into relatively longer VOTs (for example, [tʰu] *tu* 'you', [ˈsɛrpʰe] *serpe* 'snake', [ˈbɛkːʰo] *becco* 'beak'). In my sample, male speakers had longer releases in these contexts; moreover, a qualitative perception follow-up noted that participants showed different levels of awareness of the *gorgia enfatica*. Overall, the *gorgia enfatica* could be defined as a male covert-prestige feature, based on the retrieved meanings of iconized articulatory strength, rusticity (stemming from exaggeration of dialectality) and ideological firmness (stemming from both emphasis and precision). Assuming the definition of masculinity as a "bundle of stances" (Kiesling, 2001: 252), the various contextual meanings of the *gorgia enfatica* can be wrapped together in a local gender identity. Moreover, given the coherent production pattern favoring male speakers, it can be assumed that the indexical order of the socio-semiotic reinterpretations of the Florentine feature loops around masculinity

and hyperarticulation, as the former can be both the endpoint and the start (e.g., Eckert, 2008: 471) of ideological[2] linguistic processes.

It could be worth to try and see if Florentine speakers are able to extract the "male" meaning from the *gorgia enfatica* and use it in denotation. This operation has one major caveat: it tries to pattern gender as a social construct with biological sex. As I reviewed earlier (§ 1.2), the frequency code does allow for early "social" exploitation strategies of size-dependent F_0s. Nonetheless, since VOT length has no certain biological-indexical relationship with utterer's sex (Herd, 2020: 597), the differences between our object of analysis and Ohala's frequencies should be taken into account.

3. *Prepping up for the experiment*

For this first experiment on the sound-symbolic link between social and denotational meaning domains, it seems more than fitting to roll back to the basics of this line of research: forced-choice tasks (e.g., Sapir, 1929). This chapter will deal with the building blocks of this type of protocol: the ideation of an effective experimental frame, the construction of nonword stimuli *(§* 3.1), and their modality of presentation *(§* 3.2). Lastly, a brief note will cover the setting of the experiment *(§* 3.3).

3.1 Nonsense animals

Social meaning feeds on communication context (Eckert, 2012): for this reason, studies on this topic make rare use of nonwords (Munson, 2010: 173). Thus, a good balance between stimulus plausibility and distance from real lexical items (to avoid bias) seems a pivotal concern for the experiment attempted here.

For the investigation of the property of sex, two frames are seemingly viable: human and animal names. An increasing body of evidence suggests that anthroponyms manifest sound-symbolic patterns pointing at both physical and personality traits: for example, sharp-sounding names such as "Kirk" are associated with angular faces and harsh personalities (see the literature reviewed in Sidhu, Pexman, 2019). However, the search for sex patterns in proper names stumbled on the realization that these are subject to local norms of social desirability. The local relevance of names associated with strongly characterized individuals and cultural interpretations of specific sound-human attribute linkages may attenuate or completely override expected naming strategies, such as using "small" vowels (e.g., /i e/) for female anthroponyms (*ibid.*; Wang, 2022: 104). The few hints on Florentine masculinity retrieved from the above-described study are too scant to have a clear picture on this issue leading to a reasonable interpretation of the results. Moreover,

[2] Here and below (§ 5) I use this term following the acceptation promoted through the tradition of studies referenced in Eckert (2008). For example, Irvine, Gal (2000: 35) defined the "ideological aspects of linguistic differentiation" as "the ideas with which participants and observers frame their understanding of linguistic varieties and map those understandings onto people, events, and activities that are significant to them"; the same goes for individual variants. See Irvine (2022) for a related overview.

given the diversity of our setting (e.g., Nodari, Piccardi & Calamai, 2023; § 3.3), controlling for the similarity between nonwords and real anthroponyms can be seen as a noticeable hurdle. Similarly, it has been noted that animal (common) names can manifest sound-symbolic designs (Berlin, 2006), and De Carolis, Marsico & Coupé (2017: 25) endorsed their use in related experimental procedures because of the human cognitive sensitivity to animal features (see the references therein and also Barrett, 2005). Relying on an animal frame could seemingly mitigate the two above-discussed issues related to anthroponymy at the cost of creating a conceptual gap between the supposed trigger of the effect (a socio-indexical feature pertaining to human semiotics) and the endpoint of the forced choice. However, if Ohala's (1984) ethological proposal is at work when people evaluate the size of tables (as in Sapir, 1929), this should be a minor theoretical issue to confront.

In order to create Italian nonword stimuli representing animal names while controlling for psycholinguistically relevant parameters, I relied on WORDEN (Origlia, Cangemi & Cutugno, 2015), a Python interface built on the PhonItalia (Goslin, Galluzzi & Romani, 2013) knowledge base. Through WORDEN, I generated ˈCVCV nonwords while critically considering the following aspects.

– Even though the debate on the issue is far from being settled (Shinohara, Uno, 2022), I relied on Kawahara, Shinohara & Kuchimoto (2008) and assumed that the psycholinguistically prominent first segment of the nonwords was the optimal slot to trigger sound-symbolic effects. For this reason, WORDEN was set as to generate forms containing initial voiceless plosives /k t p/ (33, 11 per type), leaving the manipulation of consonant release to a subsequent phase (§ 3.2). 22 distractors with initial non-plosive segments were also analogously generated. The second consonant (whose properties were inserted as statistical control in the analysis; see § 4.4) was free to vary for the sake of enriching the stimulus battery.
– While the first vowel of the nonwords was controlled by setting it to a default /a/, vowel ending represented a major issue in stimulus construction. In fact, not only Italian marks grammatical gender through vowel ending (which may condition the perception of referent's biological sex; cf. Vigliocco, Vinson, Paganelli & Dworzynski, 2005), but previous research (Danesi, 1998) has also shown that Italian speakers assign grammatical gender to nonwords following sound-symbolic associations with relevant features of the proposed referents. Creating nonword stimuli with agrammatical vowel endings (or removing them tout court) would have implied a drastic reduction of their plausibility as Italian lexical items, hindering the access to local socio-indexical knowledge. For this reason, I decided to embrace this presumably strong grammatical bias and tried to add a balanced number of stimuli with the three Italian singular vowel endings (18 /a/, 20 /e/, 17 /o/) while controlling for their effect in the statistical analysis (§ 4.4). Examples of generated nonwords are /ˈkafa/, /ˈpave/, /ˈtatʃo/, /ˈnama/ (distractor).
– Previous research questioned the role of neighbor effects in sound symbolism (e.g., Preziosi, Coane, 2017). Given the "social" premises of the experiment presented here, having estimates of the number of very similar real Italian words

and their frequency is an invaluable asset for the evaluation of potential stimulus idiosyncrasies. Through WORDEN/PhonItalia, both phonological neighborhood density (M 10.74, SD 2.95) and log frequency (M 2.54, SD 0.51) were computed for each nonword and inserted as a control in the statistical analysis (§ 4.4).

3.2 Acoustic (and visual) stimuli

It has been argued (e.g., Bottini, Barilari & Collignon, 2019) that orthographic experience acts alongside acoustic properties in triggering sound-symbolic effects. In this work, an aural presentation of the stimuli was preferred not only for the reduction of unwanted writing bias, but also because aspirated plosives are not conventionally represented in Italian orthographic norms. Thus, in order to have a high degree of control over the experimental manipulation, aural stimuli based on the above-described (§ 3.1) nonword list were created as follows. Given the centrality of the sex parameter in this protocol, and its effects in steering the social evaluation of aural stimuli (e.g., Laur, 2008), two Florentine speakers (a 26-year-old man and a 27-year-old woman) read the nonword list in the anechoic chamber of the Scuola Normale Superiore of Pisa. Four repetitions of the 55 items list were recorded (44.1 kHz, 24 bit): two unsupervised and two guided productions. In the latter case, speakers were asked to "emit a puff of air" after initial voiceless plosives. In accordance with Kim's (1970: 107) "traditional" definition of aspiration, this instruction replicated the perceptual protocol of Piccardi (2017a) through which the social meanings attributed to the Florentine feature were originally retrieved. The production of an aspirated plosive was pinpointed through positive feedback during this training phase. The best repetitions per condition and speaker were selected in terms of overall clarity. The resulting unsupervised, unaspirated set had an overall mean VOT of 23.3 ms (12.33 SD) - [k], M 34.14, SD 11.06; [t], M 19.23, SD 8.73; [p], M 16.54, SD. 8.83. Adapting Nielsen's (2011) manipulation strategy, the bursts and the releases of these plosives were replaced in Praat (Boersma, Weenink, 2023) with the equivalent components of their aspirated counterparts from the guided productions. This aspirated set had an overall mean VOT of 127 ms (445.06 % increase from the unaspirated set overall mean; 23.41 SD) - [k^h], M 130.95, SD 28.1; [t^h], M 123.36, SD 22.96; [p^h], M 126.68, SD 17.28. The full batch of 176 recordings (((33 targets x 2 aspiration conditions) + 22 distractors) x 2 voices) had a mean duration of 499.68 ms (86.96 SD).

Forced-choice sound symbolic association tasks usually rely on visual stimuli to exemplify the experimentally relevant features of the referents (Köhler, 1930[5]). For the sake of providing participants with a visual reminder of the inquired meaning contrast, I prepared a merged picture of a blue Mars (left) and pink Venus symbol (right) on a white background while keeping an eye on the guidelines expounded in Farnand, Fairchild (2014).

3.3 A brief note on the target population

In order to run my pilot test on the sound-symbolic potential of Florentine long-lag voiceless plosives, I asked for the involvement of a senior high school class of a *liceo scientifico* in Scandicci (Florence; see Piccardi, 2022). Indeed, senior highschoolers are already peeking into the adulthood stage of their social lives (Eckert, 1997), ensuring some level of comparability with the participants in Piccardi (2017a).

Two procedures were conducted alongside the main experiment for the sake of knowing the students better while gathering a set of profiling variables for my statistical analysis. Firstly, they were individually interviewed while covering through a semi-structured topic guide several profiling topics (§ 4.4). Then, they filled in the 44-item Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, De Groot, 1990) through a 7-point Likert scale. This tool was designed to quantitatively estimate five factors pertaining to student engagement in school and learning: self-efficacy, intrinsic value, test anxiety, cognitive strategy use, and self-regulation. In my specific context, these factors could help shed some light on the effects of potential individual approaches to the completion of the experimental procedure; moreover, as Nodari (2022) pointed out, attitudes towards school can be relevant predictors of student involvement in local dialectal practices[3]. This version of the questionnaire was adapted and translated through the help of an expert Italian psychologist, and subsequently tested for its reliability (whole questionnaire: $\alpha = 0.89$; all individual factors: $\alpha \geq 0.69$).

4. *The experiment*

4.1 Participants and setting

22 senior year high school students (15 boys, 7 girls, aged 18 to 19) took part to the experiment, which was conducted in their school library in Scandicci (Florence). The room was divided into three separate activity corners for the contemporary deployment of the profiling interviews, the questionnaire completion, and the forced-choice task on sound symbolism.

4.2 Materials

The experiment makes use of the full batch of 176 aural stimuli and the background image described in § 3. These were inserted in a custom Perceval (André, Ghio,

[3] The author conducted fieldwork in two high schools in Lamezia Terme, a municipality situated in the Southern Italian Region of Calabria, in order to provide a sociophonetic account of a Regional Italian feature – allophonic aspiration targeting post-consonant and geminate voiceless plosives. Her student participants were classified in three "attitudes towards school" levels (negative, neutral, and positive), and this variable emerged as pivotal in the explanation of the aspiration patterns. For example, male speakers with negative attitudes produced the longest VOTs in a read speech task; this result was interpreted as the expression of a rough, "anti-school" male persona embracing local linguistic features as opposed to the supralocal norms promoted through formal education (*ibid.*: 231).

Cavé & Teston, 2016) script run through a Windows laptop. The stimuli were heard through a pair of Sennheiser HD 25 Light closed headphones.

4.3 Procedure

One student at a time entered the library every 20 minutes and, after signing an informed consent form, sat with me for the preliminary interview. Then, I led the student to the questionnaire corner and asked him/her to fill in the MSLQ without any time limits, while one of his/her peers took his/her place at the interview corner. Lastly, the student was accompanied to the pc corner and performed the forced-choice sound-symbolism task. Both the MSLQ and the pc experiment were completed without constant supervision. Nonetheless, I was ready to intervene at any sign of perplexity or difficulty.

The Perceval session had the following structure: firstly, instructions were textually presented, framing the experiment into a fictitious discovery of new animal species, and asking participants to match singularly presented aural stimuli with the supposed name of the male or female specimen. No time limit was coded into the experiment, but participants were invited to make their decisions using the left (for male) or right (for female) keyboard arrows as quickly as possible. Then, the 176 stimuli were presented as a sequence of randomized trials with 1.5 secs of inter-stimulus interval (i.e., a silent black screen). During each trial, one stimulus was reproduced together with the Mars/Venus (§ 3.2) image, and participants had to choose between male and female via the assigned keyboard arrows. Both individual choices and response times (RT) were recorded through Perceval. The mean duration of the Perceval sessions was 7.42 minutes (2.11 SD).

4.4 Research hypothesis, pretreatment, and analysis

The Perceval experiment tested for the hypothesis that aspirated nonwords have a higher probability of being associated with male animal specimens than non-aspirated and distractor stimuli. To this aim, 3,872 trials were recorded through the software. Given the unsupervised experimental procedure and the absence of explicit RT limits, an interquartile range criterion of outlier detection was imposed to the response dataset, ultimately pruning 289 observations (RT > 5,263 ms). The results of the remaining 3,583 trials were dummy coded (1 = M, 0 = F) and inserted as response in a generalized linear mixed effect model structure (binomial family) built in R (R Core Team, 2023; lme4 package: Bates, Maechler, Bolcher & Walker, 2015) including Participant and Item as random factors. The main "Condition" predictor (aspirated, filler, unaspirated) was evaluated together, and in interaction with several controls. Below is a comprehensive list of the tested factors, together with basic descriptive information of the variables which were not previously discussed in § 3.

– Stimulus-related: Target consonant (/k/ vs. /t/ vs. /p/ vs. filler); Vowel ending (/a/ vs. /e/ vs. /o/); Second consonant voicing (37 voiced vs. 18 voiceless in the original

nonword list; cf. D'Onofrio, 2014); Voice sex (male vs. female); Phonological neighborhood density; Log phonological neighborhood frequency.
– Participant-related: Sex (male vs. female); Regionality index score[4] (M 2.68, SD 1.82); a qualitatively assigned, two-level categorization of Attitude towards the Florentine dialect (9 positive vs. 13 nonpositive); a mean of the number of qualifications obtained by the student parents (i.e., Parent education: max. 4, min. 1, M 3.09, SD 0.65); Musical proficiency (10 proficient vs. 12 non-proficient; cf. Hallam, 2017); Pet ownership (15 owners vs. 7 non-owners); Number of languages known by the student other than Italian (M 2.27, SD 0.69); MSLQ Self-efficacy (M 4.56, SD 0.66); Intrinsic value (M 4.99, SD 0.86); Test anxiety (M 3.76, SD 1.55); Cognitive strategy use (M 5.25, SD 0.82); Self-regulation (M 4.64, SD 0.92).
– Session-related: RT (pruned dataset M in ms 1,943.66, SD 1,026.44).

Given the number of factors of interest, an incremental approach to model construction was adopted (e.g., Tordini, Galatà, Avesani & Vayra, 2018). Initially, a baseline model containing the main Condition predictor together with Participant and Item as random factors was built; since an evident pattern by Vowel ending was expected (§ 3.1), this control was immediately added to the baseline model as well. Then, the addition of the other factors as individual predictors or in interaction with Condition or the other significant controls was evaluated through χ^2 likelihood ratio tests and comparisons between Akaike and Bayesian Information Criterion scores (see Hay, Foulkes, 2016 for these three methods in binomial mixed effect model selection processes). Individual factors and interactions were retained if their addition granted significant model differences and a lowering of the relative goodness of fit scores.

4.5 Results

Overall, the "male" category was selected in 1,878/3,583 trials (52.41 %). Vowel ending was of paramount importance in determining participant responses (percentages of "male" per category: /a/, 13.32 %; /e/, 51.36 %; /o/, 94.95 %); nonetheless, once morphological gender was controlled for, two other small, but significant effects emerged: Condition (aspirated, 52.05 %; filler, 48.14 %; unaspirated, 55.56 %) and stimulus Voice sex (female, 50.28 %; male, 54.57 %). Moreover, the interaction between Vowel ending and RTs (z-scored through the scale() function out of convergence necessity) was also retained. Tab. 1 shows the final model, and Fig. 1 manifests the directionality of its interaction (sjPlot package: Lüdecke, 2023).

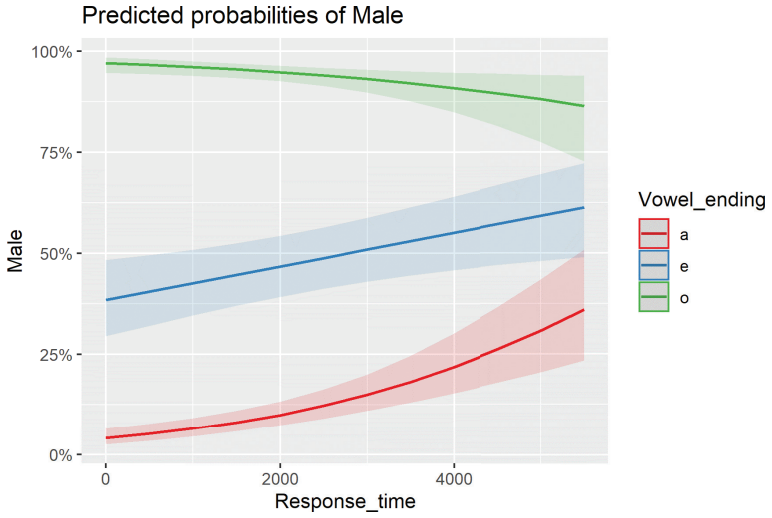
[4] The Regionality Index (Chambers, Heisler, 1999) is an estimate of ethnic affiliation. In the version used in this study, the sum of the scores attributed to the responses to four questions (i.e., student birthplace; student's parent birthplaces; student place of residence; significant relocations) constituted the individual levels of regionality. Each response ranged from 0 ("in Scandicci/Florence") to 2 ("outside the *Provincia di Firenze*") through 1 ("in the *Provincia di Firenze*"), for a maximum summed score of 8.

Table 1 - *Summary of the final model calculating the changes in log odds of responding "male" per unit change of the predictor. Recall that the other categories in the reported variables are Condition (aspirated), Voice sex (female), Vowel ending (/a/). Observations: 3,583. Random factors: Item (variance .08; SD .29); Participant (variance .23; SD .48)*

	Estimate	*Std. Error*	*z value*	*Pr(>|z|)*
(Intercept)	-2.23	0.17	-13.24	< 0.001***
Condition (filler)	-0.34	0.14	-2.51	0.012*
Condition (unaspirated)	0.29	0.12	2.35	0.019*
Voice sex (male)	0.30	0.11	2.85	0.004**
Vowel ending (/e/)	2.09	0.12	16.72	< 0.001***
Vowel ending (/o/)	5.14	0.18	28.41	< 0.001***
z Response time	0.48	0.08	5.91	< 0.001***
z Response time * Vowel ending (/e/)	-0.30	0.10	-3.04	0.002**
z Response time * Vowel ending (/o/)	-0.78	0.14	-5.43	< 0.001***

To sum up, contrary to our expectations, unaspirated nonwords had the relatively highest probabilities of being associated with male specimens, followed by aspirated and filler nonwords (in this order). Moreover, nonwords uttered by a male voice were more likely associated with a male animal specimen than nonwords uttered by a female voice. Lastly, while, overall, participants made their decisions depending on morphological gender, they tended to less morphologically congruent responses as their RT increased.

Figure 1 - *Response Time * Vowel ending interaction plot from the model shown in Tab. 1. RTs were transformed back to ms for the sake of interpretability*



5. *Discussion*

This work tried to provide a first experimental replication of an underdiscussed mechanism pertaining to Ohala's frequency code theory and to its proposed juncture between indexical and referential meaning. Specifically, I posited that Florentine speakers could be able to associate aurally presented nonword stimuli containing socially meaningful long-lag voiceless plosives with referents characterized through properties coherent to the indexical content of the variants. I expected to find an association pattern matching "manly" long VOTs with the names of fictitious male animal specimens. For the sake of guaranteeing the access to the social meaning knowledge of my participants, I decided to build my stimulus battery respecting the structure of real Italian words and included vowel endings marking grammatical gender. This operative necessity provided participants with a clear association-making strategy: responses were mostly guided by vowel ending (i.e., /o/ for "male", /a/ for "female" and /e/ for both), and any other choice pattern was rather weak and dependent on the statistical control of this stimulus feature. Nonetheless, in this chapter, I will advance an interpretation of the present results, aiming to encourage future verifications through similar endeavors.

While the null hypothesis that there is no relationship between response and experimental Condition was indeed rejected through regression modeling, effect directionality run in the opposite direction to the one which was initially advanced: aspirated nonwords had a (slightly) lower, and not higher, probability of being associated with male specimens than their unaspirated counterparts. Moreover, distractors without initial plosive segments had the lowest probability of being associated with male specimens. This result is in line with a strong cross-cultural anthroponymic trend favoring sonorants in female names (e.g., Kawahara, Katsuda & Kumagai, 2019; Wang, 2022), which may depend on associations between femininity and curviness (i.e., *maluma*-like sounds; Köhler, 1930[5]) "dating back millennia" (Sidhu, Pexman, 2015: 17). Marketing research added another piece to our explanatory attempt noting that products with names containing fricatives (vs. plosives) are perceived as more feminine (Klink, 2000). Studies searching for sound-symbolic relationships between plosive VOT and sex are rare. After a corpus analysis of Korean anthroponyms, Sullivan (2020) proposed to both English and Korean participants a name gendering experiment in which pseudo-proper names were evaluated on a scale of appropriateness from female to male. Both plosive VOTs and F_0s of the respective syllable onsets were manipulated. Long VOTs skewed Korean responses towards "male", while this feature had no effect on English associations, leading Sullivan (2020: 58) to conclude that the VOT-sex pattern is language specific. Other than on stark differences between the phonological organizations of the plosive series in the two tested languages[5], the

[5] An anonymous reviewer argued that the weak results of the Italian experiment presented here might heavily depend on phonological factors: Italian does not make distinctive use of aspiration, and aspirated voiceless plosives are absent from its phonological inventory. Indeed, Akita (2015) stressed the

Korean pattern might depend on sound-symbolic consonant alternations evoking emphasis through (phonological) aspirated plosives in word connotation (Kim, 1977), and on plausible ideological links between masculinity and exaggeration (Piccardi, 2017a: 98 and references therein). An alternative explanation could draw Sullivan's results closer to our object of inquiry. Note that her aural stimuli were built on the recordings of a middle-aged Seoul Korean man. In this variety, a sound change in progress is shifting the phonological marking of plosive aspiration from VOT to F_0, with men lagging behind women and thus showing longer release phases. Crucially, it has been argued that this social pattern is perceptually salient (Bang, Sonderegger, Kang, Clayards & Yoon, 2018: 121 and references therein). While any consideration on the perceptiveness of Sullivan's Korean participants to Seoul sociolinguistically grounded VOT distribution should be advanced with caution, voice sex may stand as an important factor in the interpretation of the Florentine results. In fact, the previous comments on the local social meaning of the inquired feature were triggered through a male voice enactment of words containing long-lag VOTs or through participant guided productions (Piccardi, 2017a), while the effects of hearing Florentine female voices using aspiration are currently unknown. As a product of hyperarticulation, long VOTs could be associated with more general female "clear" speech styles, whose frequent manifestations are explained either through sociophonetic or biological accounts (e.g., Munson, 2007: 140 and references therein). Of course, such differential triggering of social evaluations should lead not just to the general association (i.e., with the "female" category, through widespread clear speech patterns) relatively winning over the local one (i.e., with the "male" category, through Florentine masculine speech styles), but to the emergence of a statistically significant interaction between Condition and Voice sex. As this was not the case (Tab. 1)[6], further inquiries on the sound-symbolic consequences of VOT/voice sex social meaning compositionality are warranted.

Moving on to the significant controls, we find that, indeed, Voice sex had a small effect on participant's choices: nonwords uttered by a male voice were more frequently associated with male animal specimens. This result most closely comments on our research interest. Since male mammals are usually bigger than females (e.g., Lindenfors, Gittleman & Jones, 2007), and mammals are "prototypical" animals

idea that phonological typology shapes the emergence of sound-symbolic associations, and Godoy, de Souza Filho, de Souza, França & Kawahara (2020: 719) provided a short review of the literature supporting this view. The monolingual design of the present study is not suited to comment on this intriguing hypothesis, prompting further testing with other languages (§ 6).

[6] In order to double-check this point, I re-added the interaction of interest to the final model shown in Tab. 1. None of the three criteria of variable retention mentioned in § 4.4 favored the more complex model: the $\chi 2$ likelihood ratio test between the Tab. 1 and the Tab. 1 plus the interaction between Condition and Voice sex models did not reach statistical significance, and the latter showed higher AIC and BIC scores, with a difference of more than 2/less than 4 and more than 10, respectively (Raftery, 1995; Burnham, Anderson, 2004). When computed in addition to the Tab. 1 formula, the interaction had the following output: Condition (filler) * Voice sex (male): est. -.2, std. error .27, p = .46; Condition (unaspirated) * Voice sex (male): est. .04, std. error .25, p = .87.

from a student viewpoint (Trowbridge, Mintzes, 1988), the results pertaining to Voice sex may be considered as a roundabout replication of Ohala's F_0-to-size mapping: the low-F_0 human male voice could have been associated with implicitly mammal, generally bigger male animal specimens. However, previous research (Rojczyk, 2011) failed to induce referent size sound-symbolic associations through F_0 manipulation alone. Thus, an alternative explanation of our Voice sex effect could imply a more direct sound-symbolic mapping involving the second frequency code layer of biologically determined indexicality, i.e., the one concerning human sexual dimorphism (§ 1.2). In other words, participants could have extracted general features (including different mean F_0s) of human male vs. female voice and repurposed them to identify living beings in a sexually coherent fashion. While the exact triggers (i.e., the co-variating elements characterizing gendered voice, which may also go beyond the biological into the social realm) of this matching strategy cannot be ascertained through the current experiment, this result highlights that, indeed, speakers can bring out indexical information from acoustic stimuli and predictably use it to depict features of fictitious referents.

Taken together, these two results highlight the role played by co-occurrence consistency (i.e., frequency) between specific indexical meanings and their acoustic means of expression in order to successfully prompt coherent transfers to denotational sound-symbolic mechanisms: the retrieved patterns relatively favored basic socio-semiotic gender characterization of hyperarticulated speech over local social meaning, and general properties of human voice sex. Statistical co-occurrence is indeed one of the main proposed explanatory factors for sound symbolism. Sidhu, Pexman (2018: 1626-1627, 1631) noted that frequency-to-size mappings can be easily framed within this principle, which is also the basis of Ohala's (1984) frequency code. In Piccardi (2017b) I argued that, seeking for a model grounded in the encoding of both eventive frequencies and indexical properties, sound symbolic theory could benefit from the adoption of an exemplarist mindset. Exemplar theory posits that humans are capable of storing memories of individual communicative events in probabilistic structures, retaining fine-grained phonetic details and contextual information (e.g., Pierrehumbert, 2006). In this framework, previous research suggested that contextually evoking specific concepts can activate related language exemplars which, in turn, can steer perceptions (Hurring, Hay, Drager, Podlubny, Manhire & Ellis, 2022) and productions (Sanchez, Hay & Nilson, 2015) towards conceptually coherent targets. Crucially, Hurring *et al.* (2022) recently provided tentative confirmation to the idea that, in order to take place, these contextual priming effects may need a consistent prior exposure of the subject to the linguistic elements linked to the evoked concept. Given this theoretical background, Ohala's (1984) explanation of indexical-to-denotational meaning transfers (§ 1.2) can be rephrased as follows. The biological motivation of F_0 variation in the animal world creates a highly populated cluster of eventive traces encoding information about both vocalization frequency and contextual size (of the utterer). Encountering visual cues denoted by a size-related concept (i.e., "bigness")

activates this clustering, together with its attached acoustic detail. Note that, in the absence of arbitrary word-to-meaning associations (which are roughly minimized using nonwords in sound-symbolic experiments), the speaker's reference exemplar distribution is vastly based on prior indexical experiences. Thus, when prompted to perform a naming (or forced-choice) task, the speaker may be induced to resort to the main acoustic information attached to the activated concept of size, which is originally of indexical nature. While previous research seldom felt the need of making sound symbolism talk to exemplar theory (but see Auracher, Scharinger & Menninghaus, 2019: 12), it may be argued that this framework could lead to a more organic treatment of indexicality and statistical co-occurrence in this line of work[7].

Last in proximity to our research interest but first in magnitude, the effects of Vowel ending on the forced choices provide confirmation to the idea that Italian grammatical gender can affect word meaning, at least when referring to sexed entities (Vigliocco *et al.*, 2005). While a full recount on this topic is well outside the scope of this paper (see Samuel, Cole & Eacott, 2019 for a recent review), it should be noted that /e/ endings, which are "opaque" from the viewpoint of grammatical gender, triggered the most uncertain association pattern and the slowest responses[8] (see the literature gathered in Padovani, Calandra-Buonaura, Cacciari, Benuzzi & Nichelli, 2005: 301), supporting the primacy of grammatical markers in the observed decision process. The interaction between Vowel ending and RT may imply that longer evidence accumulation processes (e.g., Ratcliff, Smith, Brown & McKoon, 2016) are required to reach morphologically unintuitive decisions. However, this result cannot provide any further detail on the type of knowledge added to grammatical gender in the slow responses, nor on its exact source.

6. *Conclusions*

This work finds its place in an ongoing revival of research on sound symbolism. In this milieu, third-wave sociolinguists are enthusiastically considering sound symbolism as an important indexical resource through an endorsement of Ohala's (1984) frequency code, legitimating its presence in their disciplinary paradigm and an incipient acceptation of meaning which encompasses the social and the

[7] This is not by any means to say that exemplar theory is the only model suited to deal with sound symbolism. For example, Kawahara (2021) recently suggested that Maximum Entropy Harmonic Grammar can account for sound-symbolic patterns. Moreover, indeed, other theoretical perspectives in phonology have frequency and variation at heart (Coetzee, Pater, 2011; Alderete, Finley, 2023). Here, exemplar theory is mentioned in reference to the specific issue at hand, which concerns indexical knowledge and might be dependent on the relative frequency of events in which both acoustic and contextual features are concurrently relevant. As mentioned above, the proposal of a probabilistically structured "unified mental architecture" (Pierrehumbert, 2006: 520) embedding these heterogeneous aspects of communicative experience is central to exemplar theory in linguistics (see also Foulkes, Docherty, 2006).

[8] This was verified by fitting a restricted maximum likelihood linear mixed effect model with ms RT as dependent variable, Vowel ending as predictor and Participant and Item as random effects. /e/, and not /o/, (from a baseline /a/) was significant (est. 406.82, std. error 40.33, $p < .001$).

denotational alike (D'Onofrio, Eckert, 2021). Nonetheless, as per admission of Ohala (1984) himself, the frequency code is not well-equipped to justify potential sound-symbolic meaning leakages between these two domains. Thus, I attempted to experimentally verify the supposed human capability of extracting meaning features at the indexical level and repurposing them in denotation by submitting a forced-choice task to a senior high school class in Scandicci (Florence).

The experiment only managed to glimpse a very feeble light at the end of this disciplinary tunnel. For the sake of providing naturalistic nonwords to the social evaluation of Italian participants, I committed to the creation of stimuli with vowel endings marking grammatical gender. This choice had a huge impact on the experimental results, as participants made evident use of them in their decision processes. While, of course, the removal of such bias does not guarantee the emergence of alternative patterns (or the strengthening of the ones commented here), future studies should be primarily concerned with testing Ohala's assumptions in a language without noun gender markers steering participant associations. English is the obvious solution to this conundrum, and not only because of its lack of noun grammatical gender: this language is the homeland of third-wave sociolinguistics and indexical field representations (Eckert, 2008; see § 2), so that similar research can be supported by solid acquisitions on the social meaning nuances of linguistic variables potentially orienting denotational sound symbolism.

Other than vowel ending, two other factors had a very small impact on the experimental results. Aspiration did condition participant reactions, but in an unexpected way: in fact, long-lag nonwords had a little higher chance of being associated with female, and not male, animal specimens. This result was tentatively interpreted as a consequence of relatively prevailing, more general (and thus, less locally bound), patterns between female speakers and clear, hyperarticulated speech, which were probably evoked through the presence of a female voice producing long VOTs in the stimulus battery. Perceived voice sex manifested a less ambiguous pattern: nonwords uttered by a male voice were more likely to be associated with a male specimen. This specific effect gives some plausibility to the idea that, at some level, participants were able to extract the indexical property of sex from the acoustic stimuli and recycle it in the denotation of unknown living entities.

These two almost negligible effects served to highlight a current partial limitation of sound-symbolic research. In the mid-seventies, Hecht (1976) rightly promoted a more articulate modelling of respondent social characteristics in sound symbolism research. Now, sociolinguistics (or sociophonetics) may be called to answer to novel variationist issues: how does frequency of co-occurrence between the two main components of a sound symbolic association (e.g., a sound and an indexical meaning) affect its probabilistic manifestations (cf. Sidhu, Pexman, 2018: 1627)? How can we account for the potential effects of experiment contextual characteristics (e.g., the features of the stimulus utterers, such as sex, age, speech variety, but also environmental conditioning) in protocols investigating sound symbolism? The adoption of

frequency-grounded, contextually informed exemplar models might be the next step for the consolidation of the sociolinguistic tradition on the matter at hand.

Acknowledgements

This paper is the result of a Doctoral spin-off project held at the University of Pisa. I wish to thank my Italian advisors, Giovanna Marotta and Silvia Calamai, for their invaluable guidance. My deepest gratitude goes also to the involved personnel of the Russel Newton high school in Scandicci, namely Anna Maria Addabbo, Susanna Bino, and Luca Gargiulo, as well as to the V C students.

References

Acton, E.K. (2021). Pragmatics and the Third wave: The Social Meaning of Definites. In Hall-Lew, L., Moore, E. & Podesva, R.J. (a cura di), *Social Meaning and Linguistic Variation. Theorizing the Third Wave*. Cambridge: Cambridge University Press, 105-126.

Akita, K. (2015). Sound Symbolism. In Östman, J.-O., Verschueren, J. (a cura di), *Handbook of Pragmatics. 2015 Installment*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/hop.19.sou1

Akita, K. (2021). Phonation Types Matter in Sound Symbolism. In *Cognitive Science*, 45(5), 1-14. https://doi.org/10.1111/cogs.12982

Alderete, J., Finley, S. (2023). Probabilistic Phonology: A Review of Theoretical Perspectives, Applications, and Problems. In *Language and Linguistics*, 24(4), 565-610.

Alderete, J., Kochetov, A. (2017). Integrating Sound Symbolism with Core Grammar: The Case of Expressive Palatalization. In *Language*, 93(4), 731-766.

André, C., Ghio, A., Cavé, C. & Teston, B. (2016). PERCEVAL: PERCeption EVALuation Auditive & Visuelle. [Computer program] Version 4.0.0. https://lpldev.lpl-aix.fr/perceval/

Auracher, J., Scharinger, M. & Menninghaus, W. (2019). Contiguity-Based Sound Iconicity: The Meaning of Words Resonates with Phonetic Properties of their Immediate Verbal Contexts. In *PLoS ONE*, 14(5), 1-21. https://doi.org/10.1371/journal.pone.0216930

Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M. & Yoon, T.-J. (2018). The Emergence, Progress, and Impact of Sound Change in Progress in Seoul Korean: Implications for Mechanisms of Tonogenesis. In *Journal of Phonetics*, 66, 120-144.

Barrett, H.C. (2005). Cognitive Development and the Understanding of Animal Behavior. In Ellis, B.J., Bjorklund, D.F. (Eds.), *Origins of the Social Mind. Evolutionary Psychology and Child Development*. New York / London: The Guildford Press, 438-467.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. In *Journal of Statistical Software*, 67(1), 1-48.

Berlin, B. (2006). The First Congress of Ethnozoological Nomenclature. In *The Journal of the Royal Anthropological Institution*, 12(s1), 23-44.

Besnier, N. (1990). Language and Affect. In *Annual Review of Anthropology*, 19, 419-451.

Boersma, P., Weenink, D. (2023). PRAAT: Doing Phonetics by Computer. [Computer program] Version 6.3.10. https://www.praat.org/

Bottini, R., Barilari, M. & Collignon, O. (2019). Sound Symbolism in Sighted and Blind. The Role of Vision and Orthography in Sound-Shape Correspondences. In *Cognition*, 185, 62-70.

Bucholtz, M., Hall, K. (2016). Embodied Sociolinguistics. In Coupland, N. (a cura di), *Sociolinguistics: Theoretical Debates*. Cambridge: Cambridge University Press, 173-198.

Burnham, K.P., Anderson, D.R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. In *Sociological Methods & Research*, 33(2), 261-304.

Castellani, A. (1960). Precisazioni sulla gorgia toscana. In *Boletim de Filología*, 19, 242-262.

Chambers, J.K., Heisler, T. (1999). Dialect Topography of Quebéc City English. In *Canadian Journal of Linguistics*, 44(1), 23-48.

Coetzee, A.W., Pater, J. (2011). The Place of Variation in Phonological Theory. In Goldsmith, J.A., Riggle, J. & Yu, A.C.L. (Eds.), *The Handbook of Phonological Theory*. Second Edition. Oxford: Blackwell-Wiley, 401-431.

D'Onofrio, A. (2014). Phonetic Detail and Dimensionality in Sound-shape Correspondences: Refining the Bouba-Kiki Paradigm. In *Language and Speech*, 57(3), 367-393.

D'Onofrio, A., Eckert, P. (2021). Affect and Iconicity in Phonological Variation. In *Language in Society*, 50(1), 29-51.

Danesi, M. (1998). Gender Assignment, Markedness, and Indexicality: Results of a Pilot Study. In *Semiotica*, 121(3-4), 213-240.

De Carolis, L., Marsico, E. & Coupé, C. (2017). Evolutionary Roots of Sound Symbolism. Association Tasks of Animal Properties with Phonetic Features. In *Language & Communication*, 54, 21-35.

Dressler, W.U., Merlini Barbaresi, L. (1992). Italian Diminutives as Non-Prototypical Word Formation. In Tonelli, L., Dressler, W.U. (a cura di), *Natural Morphology-Perspectives for the Nineties*. Padua: Unipress, 21-30.

Drummond, R., Schleef, E. (2016). Identity in Variationist Sociolinguistics. In Preece, S. (Ed.), *The Routledge Handbook of Language and Identity*. London: Routledge, 50-65.

Eckert, P. (1997). Age as a Sociolinguistic Variable. In Coulmas, F. (a cura di), *The Handbook of Sociolinguistics*. Oxford/Cambridge: Blackwell, 151-167.

Eckert, P. (2008). Variation and the Indexical Field. In *Journal of Sociolinguistics*, 12(4), 453-476.

Eckert, P. (2010). Affect, Sound Symbolism, and Variation. In *University of Pennsylvania Working Papers in Linguistics*, 15(2), 70-80.

Eckert, P. (2012). Three Waves of Variation Study. In *Annual Review of Anthropology*, 41, 87-100.

Eckert, P. (2019). The Limits of Meaning: Social Indexicality, Variation, and the Cline of Interiority. In *Language*, 95(4), 751-776.

Eckert, P., Labov, W. (2017). Phonetics, Phonology and Social Meaning. In *Journal of Sociolinguistics*, 21(4), 467-496.

Ekström, A.G. (2022). What's Next for Size-Sound Symbolism? In *Frontiers in Language Sciences*, 1, 1-7. https://doi.org/10.3389/flang.2022.1046637

Farnand, S.P., Fairchild, M.D. (2014). Designing Pictorial Stimuli for Perceptual Experiments. In *Applied Optics*, 53(3), 72-82.

Foulkes, P. (2010). Exploring Social-Indexical Knowledge: A Long Past but a Short History. In *Laboratory Phonology*, 1(1), 5-39.

Foulkes, P., Docherty, G. (2006). The Social Life of Phonetics and Phonology. In *Journal of Phonetics*, 34(4), 409-438.

Francis, A.L., Ciocca, V. & Yu, J.M.C. (2003). Accuracy and Variability of Acoustic Measures of Voicing Onset. In *The Journal of the Acoustical Society of America*, 113(2), 1025-1032.

Gafter, R.J. (2021). Reconciling Seemingly Conflicting Social Meanings. In Hall-Lew, L., Moore, E. & Podesva, R.J. (a cura di), *Social Meaning and Linguistic Variation. Theorizing the Third Wave*. Cambridge: Cambridge University Press, 222-242.

Glanzberg, M. (2021). Chomsky on Semantics. In Allott, N., Lohndal, T. & Rey, G. (Eds.), *A Companion to Chomsky*. Hoboken: Wiley-Blackwell, 416-432.

Godoy, M.C., de Souza Filho, N.S., de Souza, J.G., França, H.A. & Kawahara, S. (2020). Gotta Name'Em All: An Experimental Study on the Sound Symbolism of Pokémon Names in Brazilian Portuguese. *In Journal of Psycholinguistic Research*, 49(5), 717-740.

Gordon, M., Heath, J. (1998), Sex, Sound Symbolism, and Sociolinguistics. In *Current Anthropology*, 39(4), 421-449.

Goslin, J., Galluzzi, C. & Romani, C. (2013). Phonitalia: A Phonological Lexicon for Italian. In *Behavior Research Methods*, 46(3), 872-886.

Hallam, S. (2017). The Impact of Making Music on Aural Perception and Language Skills: A Research Synthesis. In *London Review of Education*, 15(3), 388-406.

Hannahs, S.J., Bosch, A.R.K. (2018). The Study of Phonology in the 21st Century. Overview and Introduction to The Routledge Handbook of Phonological Theory. In Hannahs, S.J., Bosch, A.R.K. (a cura di), *The Routledge Handbook of Phonological Theory*. New York: Routledge, 1-10.

Hay, J., Foulkes, P. (2016). The Evolution of Medial /t/ Over Real and Remembered Time. In *Language*, 92(2), 298-330.

Heath, J. (2003). Review to Principles of Linguistic Change. Volume 2: Social Factors by William Labov. In *Anthropological Linguistics*, 45(4), 466-472.

Hecht, M.L. (1976). More Semantics of Repetition: A Sociolinguistic Investigation of Phonetic Symbolism. In *Communication Quarterly*, 24(3), 33-40.

Herd, W. (2020). Sociophonetic Voice Onset Time Variation in Mississippi English. In *The Journal of the Acoustical Society of America*, 147(1), 596-605.

Hurring, G., Hay, J., Drager, K., Podlubny, R., Manhire, L. & Ellis, A. (2022). Social Priming in Speech Perception: Revisiting Kangaroo/Kiwi Priming in New Zealand English. In *Brain Sciences*, 12, 1-29. https://doi.org/10.3390/brainsci12060684

Irvine, J.T. (2022). Revisiting Theory and Method in Language Ideology Research. In *Linguistic Anthropology*, 32(1), 222-236.

Irvine, J.T., Gal, S. (2000). Language Ideology and Linguistic Differentiation. In Kroskrity, P. (Ed.), *Regimes of Language: Ideologies, Polities, and Identities*. Santa Fe: School of American Research Press, 35-83.

Johnstone, B. (2010). Locating Language in Identity. In Llamas, C., Watt, D. (a cura di), *Language and Identities*. Edinburgh: University Press, 29-36.

Jurafsky, D. (1996). Universal Tendencies in the Semantics of the Diminutive. In *Language*, 72(3), 533-578.

Karthikeyan, S., Puts, D.A., Aung, T., Link, J.K., Rosenfield, K., Mackiel, A., Casey, A., Marks, K., Cristo, M., Patel, J., Santos, A. & Geher, G. (2023). Articulatory Effects on Perceptions of Men's Status and Attractiveness. In *Scientific Reports*, 13, 1-11. https://doi.org/10.1038/s41598-023-29173-z

Kawahara, S. (2020). Sound Symbolism and Theoretical Phonology. In *Language and Linguistic Compass*, 14(8), 1-17. https://doi.org/10.1111/lnc3.12372

Kawahara, S. (2021). Testing MaxEnt with Sound Symbolism: A Stripy Wug-Shaped Curve in Japanese Pokémon Names. In *Language*, 97(4), 341-359.

Kawahara, S., Katsuda, H. & Kumagai, G. (2019). Accounting for the Stochastic Nature of Sound Symbolism using Maximum Entropy Model. In *Open Linguistics*, 5(1), 109-120.

Kawahara, S., Shinohara, K. & Uchimoto, Y. (2008). A Positional Effect in Sound Symbolism: An Experimental Study. In *Proceedings of the Japan Cognitive Linguistics Association*, 8, 417-427.

Kiesling, S.F. (2001). "Now I Gotta Watch What I Say": Shifting Constructions of Masculinity in Discourse. In *Journal of Linguistic Anthropology*, 11(2), 250-273.

Kim, C. (1970). A Theory of Aspiration. In *Phonetica*, 21(2), 107-116.

Kim, K.-O. (1977). Sound Symbolism in Korean. In *Journal of Linguistics*, 13(1), 67-75.

Klink, R. (2000). Creating Brand Names With Meaning: The Use of Sound Symbolism. In *Marketing Letters*, 11(1), 5-20.

Köhler, W. (1930[5]). *Gestalt Psychology*. New York: Liveright.

Krivochen, D.G., Lacková, L. (forthcoming). Iconicity and Generative Grammar. Manuscript, University of Oxford / Palacký University.

Labov, W. (2001). *Principles of Linguistic Change. Social Factors*. Malden/Oxford: Blackwell.

Laur, E. (2008). *Contribution à l'étude des perceptions linguistiques. La méthodologie des faux-couples revisitée*. Montréal: Office québécois de la langue française.

Lee, N.H. (2020). Style Variation in the Second Formant. What Does it Mean to be "Refined" in Baba Malay? In *Language Ecology*, 4(1), 115-130.

Levon, E., Maegaard, M. & Pharao, N. (2017). Introduction: Tracing the Origin of /s/ Variation. In *Linguistics*, 55(5), 979-992.

Lindenfors, P., Gittleman, J.L. & Jones, K.E. (2007). Sexual Size Dimorphism in Mammals. In Blackenhorn, W.V., Fairbairn, D.J. & Székely, T. (a cura di), *Sex, Size and Gender Roles: Evolutionary Studies of Sexual Size Dimorphism*. Oxford: Oxford University Press, 280-315.

Lisker, L., Abramson, A.S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. In *Word*, 20(3), 527-565.

Lüdecke, D. (2023). sjPlot: Data Visualization for Statistics in Social Science. [R package] version 2.8.14, retrieved May, 20 2023 from https://CRAN.R-project.org/package=sjPlot

Magnus, M. (2013). A History of Sound Symbolism. In Allan, K. (a cura di), *The Oxford Handbook of the History of Linguistics*. Oxford: Oxford University Press, 191-208.

Marotta, G. (2008). Lenition in Tuscan Italian (Gorgia Toscana). In Brandão de Carvalho, J., Scheer, T. & Ségéral, P. (a cura di), *Lenition and Fortition*, Berlin: Mouton de Gruyter, 235-271.

Munson, B. (2007). The Acoustic Correlates of Perceived Masculinity, Perceived Femininity, and Perceived Sexual Orientation. In *Language and Speech*, 50(1), 125-142.

Munson, B. (2010). Levels of Phonological Abstraction and Knowledge of Socially Motivated Speech-Sound Variation: A Review, a Proposal, and a Commentary on the Papers by Clopper, Pierrehumbert, and Tamati; Drager; Foulkes; Mack; and Smith, Hall, and Munson. In *Laboratory Phonology*, 1(1), 157-177.

Nielsen, A.K., Dingemanse, M. (2021). Iconicity in Word Learning and Beyond: A Critical Review. In *Language and Speech*, 64(1), 52-72.

Nielsen, K. (2011). Specificity and Abstractness of VOT Imitation. In *Journal of Phonetics*, 39(2), 132-142.

Nobile, L. (2019). Introduction. Sound Symbolism in the Age of Digital Orality. A Perspective on Language Beyond 'Nature' and 'Culture'. In *Significances (Signifying)*, 3(1), XXXVI-LXVIII.

Nodari, R. (2022). *L'identità linguistica regionale degli adolescenti. Aspirazione delle occlusive sorde in Calabria e percezione della varietà locale*. Roma: Aracne.

Nodari, R., Piccardi, D. & Calamai, S. (2023). Lingue conosciute, parlate, percepite. Il punto di vista degli studenti in Toscana. In *EL.LE*, 12(3), 385-431.

Ochs, E., Schieffelin, B. (1989). Language has a Heart. In *Text*, 9(1), 7-25.

Ohala, J.J. (1984). An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice. In *Phonetica*, 41(1), 1-16.

Origlia, A., Cangemi, F. & Cutugno, F. (2015). WORDEN: a PYTHON Interface to Automatic (Non)word Generation. In Vayra, M., Avesani, C. & Tamburini, F. (a cura di), *Language Acquisition and Language Loss. Acquisition, Change and Disorders of the Language Sound Structure*. *Studi AISV 1*. Milan: Officinaventuno, 459-470.

Padovani, R., Calandra-Buonaura, G., Cacciari, C., Benuzzi, F. & Nichelli, P. (2005). Grammatical Gender in the Brain: Evidence from an fMRI Study on Italian. In *Brain Research Bulletin*, 65(4), 301-308.

Piattelli-Palmarini, M., Cecchetto, C. (1997). The Problem of Meaning in Generative Grammar. In Mandell, C., McCabe, A. (a cura di), *The Problem of Meaning: Behavioral and Cognitive Perspectives*. Amsterdam: Elsevier, 415-469.

Piccardi, D. (2017a). Sociophonetic Factors of Speakers' Sex Differences in Voice Onset Time: A Florentine Case Study. In Bertini, C., Celata, C., Lenoci, G., Meluzzi, C. & Ricci, I. (a cura di), *Social and Biological Factors in Speech Variation. Studi AISV 3*. Milan: Officinaventuno, 83-106.

Piccardi, D. (2017b). The Social Life of Sound Symbolism. Analysing the Relationship Between Indexical and Referential Meaning. Paper presented at CODISCO X, September 29, 2017.

Piccardi, D. (2021). *Rural Sociophonetics. Dialectology Meets Usage-based Theory in the Village of Antona (Massa, Italy)*. Alessandria: Edizioni dell'Orso.

Piccardi, D. (2022). Teaching Sound Symbolism through Japanese Pop Culture. A Response to Kawahara (2018). In *Italiano LinguaDue*, 14(2), 724-758.

Pierrehumbert, J.B. (2006). The Next Toolkit. In *Journal of Phonetics*, 34(4), 516-530.

Pintrich, P.R., De Groot, E.V. (1990). Motivational and Self-regulated Learning Components of Classroom Academic Performance. In *Journal of Educational Psychology*, 82(1), 33-40.

Preziosi, M.A., Coane, J.H. (2017). Remembering that Big Things Sound Big: Sound Symbolism and Associative Memory. In *Cognitive Research: Principles and Implications*, 2, 1-21. https://doi.org/10.1186/s41235-016-0047-y

R Core Team (2023). R: A Language and Environment for Statistical Computing. [Computer program] Version 4.3.0. https://www.r-project.org/

Raftery, A.E. (1995). Bayesian Model Selection in Social Research. In *Sociological Methodology*, 25, 111-163.

Ratcliff, R., Smith, P.L., Brown, S.D. & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. In *Trends in Cognitive Science*, 20(4), 260-281.

Rojczyk, A. (2011). Sound Symbolism in Vowels: Vowel Quality, Duration and Pitch in Sound-to-Size Correspondence. In *Poznań Studies in Contemporary Linguistics*, 47(3), 602-615.

Samarin, W.I. (1978). Linguistic Adaptation to Speech Function. In McCormack, W., Wurm, S.A. (a cura di), *Approaches to Language: Anthropological Issues*. The Hague: Mouton, 559-614.

Samuel, S., Cole, G. & Eacott, M.J. (2019). Grammatical Gender and Linguistic Relativity: A Systematic Review. In *Psychonomic Bulletin & Review*, 26(6), 1767-1786.

Sanchez, K., Hay, J. & Nilson, E. (2015). Contextual Activation of Australia Can Affect New Zealanders' Vowel Productions. In *Journal of Phonetics*, 48, 76-95.

Sapir, E. (1929). A Study in Phonetic Symbolism. In *Journal of Experimental Psychology*, 12(3), 225-239.

Schertz, J. (2013). Exaggeration of Featural Contrasts in Clarifications of Misheard Speech in English. In *Journal of Phonetics*, 41(3-4), 249-263.

Scobbie, J.M. (2006). Flexibility in the Face of Incompatible English VOT Systems. In Goldstein, L., Whalen, D.H. & Best, C.T. (a cura di), *Papers in Laboratory Phonology 8: Varieties of Phonological Competence*. Berlin: Mouton de Gruyter, 367-392.

Shinohara, K., Uno, R. (2022). Exploring the Positional Effects in Sound Symbolism. The Case of Hardness Judgments by English and Japanese Speakers. In *Languages*, 7(3), 1-11. https://doi.org/10.3390/languages7030179

Sidhu, D.M., Pexman, P.M. (2015). What's in a Name? Sound Symbolism and Gender in First Names. In *PLoS ONE*, 10(5), 1-22. https://doi.org/10.1371/journal.pone.0126809

Sidhu, D.M., Pexman, P.M. (2018). Five Mechanisms of Sound Symbolic Association. In *Psychonomic Bulletin & Review*, 25(5), 1619-1643.

Sidhu, D.M., Pexman, P.M. (2019). The Sound Symbolism of Names. In *Current Directions in Psychological Science*, 28(4) 398-402.

Silverstein, M. (1994). Relative Motivation in Denotational and Indexical Sound Symbolism of Wasco-Wishram Chinookan. In Hinton, L., Nichols, J. & Ohala, J.J. (a cura di), *Sound Symbolism*. Cambridge: University Press, 40-60.

Sullivan, L. (2020.) Universal and Cross-linguistic Gender-based Sound Symbolism in Korean Given Names. Generals Paper 2. University of Toronto, Toronto.

Svantesson, J-O. (2017). Sound Symbolism: The Role of Word Sound in Meaning. In *WIREs Cognitive Science*, 8(5), 1-12. https://doi.org/10.1002/wcs.1441

Tordini, O., Galatà, V., Avesani, C. & Vayra, M. (2018). Sound Maintenance and Change: Exploring Inter-language Phonetic Influence in First-generation Italo-Australian Immigrants. In Vietti, A., Spreafico, L., Mereu, D. & Galatà, V. (a cura di), *Speech in the Natural Context. Studi AISV 4*. Milan: Officinaventuno, 77-98.

Trowbridge J.E., Mintzes, J.J. (1988). Alternative Conceptions in Animal Classification: A Cross-Age Study. In *Journal of Research in Science Teaching*, 25(7), 547-571.

Usnadze, D. (1924). Ein experimenteller Beitrag zum Problem der psychologischen

Vigliocco, G., Vinson, D.P., Paganelli, F. & Dworzynski, K. (2005). Grammatical Gender Effects on Cognition: Implications for Language Learning and Language Use. In *Journal of Experimental Psychology: General*, 134(4), 501-520.

Wang, X. (2022). Recherches sur le symbolisme phonétique des noms propres. In *Signifiances (Signifying)*, 6(1), 98-120.

Winter, B., Oh, G.E., Hübscher, I., Idemaru, K., Brown, L., Prieto, P. & Grawunder, S. (2021). Rethinking the Frequency Code: A Meta-Analytic Review of the Role of Acoustic Body Size in Communicative Phenomena. In *Philosophical Transactions B*, 376, 1-13. https://doi.org/10.1098/rstb.2020.0400

Woschitz, J. (2019). Language in and out of Society: Converging Critiques of the Labovian Paradigm. In *Language & Communication*, 64, 53-67.

Zimman, L. (2023). Language, Gender, Race, and Sexuality: Intersectional Perspectives. In Duranti, A., George, R. & Riner, R.C. (a cura di), *A New Companion to Linguistic Anthropology*. Hoboken/Chichester: Wiley Blackwell, 525-541.