Lexical Negotiations

0. Abstract

The use of lexical signs like 'knowledge' has consequences. Not only do they have direct psychological resonances, but people ascribe beliefs and act based on their semantics. This paper proposes that such consequences are up for negotiation, and introduces a formal framework from financial theory to suggest constraints on those negotiations and implications of those constraints. The upshot is that changing language will be easier sometimes than others, and philosophers' projects of linguistic change should be aware of those conditions.

1. The Problem of Lexical Negotiation

Words are often considered as either the results of specific sounds or marks or as yielding a certain syntax and semantics. Dictionaries connect these characteristics, listing words in lexical order and giving their pronunciation, part of speech, and definition. None of these characteristics are immutable, however: orthography, paleography, pronunciation, parts of speech, and semantics all vary across time and context. "Fyre" now suggests an infamous festival, not a variant spelling of "fire," and the verb "to fire" more readily suggests ending employment than setting alight. Moreover, these characteristics of words are products of intentionality. Agents make particular sounds or marks with the intention of educing a certain word, and use a certain word with the intention of educing a certain grammar and meaning. Nonetheless these aspects of words cannot be merely matters of individual authorial intentionality because they are subject to judgments of error: words can be misspelled, mispronounced, and used in ways that are grammatically incorrect, and (as frequently discussed among philosophers) words do not mean whatever their users intend (Burge, 1979).

Negotiations over when to use certain lexical signs are common in both academic philosophy and every day speech. Philosophers argue about when to use 'ground,' 'supervenience,' etc. (Plunkett, 2015), and some like David Chalmers (2011) think that these 'verbal disputes' make up much of published academic philosophy. Everyday conversations involve coming to agreement about both loaded terms like 'torture' or 'assault rifle' and banal ones like 'loafer' (Hansen, forthcoming). In ameliorative projects like Sally Haslanger (2000)'s, the goal is for philosophers to encourage a broader social revision in the use of signs like 'woman.'

As Nat Hansen (forthcoming) outlines, such metalinguistic moves can be conceptual/semantic or merely regard usage, involve explicit or merely implicit dispute/negotiation, and intend a change merely for a particular conversation or for broader social norms, which is why philosophers have used so many different terms to describe them. Further, these disputes can extend not only to the use of words, but to what orthography, paleography, and pronunciation should be accepted as indicating a particular word (Miller, 2020). I will call all of these 'lexical negotiations' since they all involve the acceptance or rejection of a speaker's desire for a lexical sign to have certain linguistic consequences. One immediate payoff of considering these diverse meta-discourse strategies under the rubric of negotiation is that we can recognize that all of their results are dependent on the relative power of the negotiators, and hence raise the specter of hermeneutical injustice (Fricker, 2006), but I would like to explore another payoff in this paper.

Optimists about the kinds of philosophical projects described above (e.g. Plunkett & Sundell, 2013) think that such metalinguistic negotiations can be successful and productive. Pessimists like Herman Cappelen (2018), by contrast, are dubious that any of these projects can succeed because of the degree of control they presume over the consequences of our usage of lexical signs. Cappelen levels two basic arguments for this pessimistic view. First, he notes that historical changes in the linguistic consequences of lexical signs are slow and often driven by relatively unconscious factors. Second, Cappelen charges that broad

philosophical acceptance of semantic externalism should be taken to suggest that changes in the linguistic consequences of lexical signs are driven by historical and physical causes outside our intentional control. Steffen Koch (2021) has responded that such considerations only imply that control is 'collective' and 'long-range.' The empirical facts and our philosophical understanding of semantics require that achieving changes in the consequences of lexical signs is not generally easy, precisely because it 'requires others to play along' but this does not mean that we utterly lack control. This might be considered a merely verbal dispute over 'control' or a pragmatic dispute over the value of pursuing projects where the degree of control is low, but it also suggests a need to understand why linguistic changes *are* sometimes more rapid and intentional and other times not. After all, real-life lexical negotiation programs often exhibit complex swings between different levels of control.

> **Commented [RM1]:** Summarize the argument.

An instructive example is the emerging distinction between 'equality' (understood as equality of distributed resources) and 'equity' (understood as fair outcomes), as represented in an extremely popular image of children standing on boxes to watch a baseball game.[1] 'Equity' and 'equality' are cognates with similar meanings, and while their historical usage supports a distinction between equality-as-sameness and equity-as-fairness, this does not map onto a distinction between inputs and outcomes (Espinoza, 2007; O'Conner & Kellerman, 2019). In fact the original 2012 version of the image referenced two different notions of equality—later versions labeled the panel with equal outcomes as 'fair' or 'just,' and only then was the 'equity' label applied (Froehle, 2020). This viral evolution seems like selection on near-random variation outside of the creator's control, justifying Cappelen's pessimism, but these changes all happened within a few months of the image's creation. The image didn't become pervasive until it was professionally redrawn by Angus Maguire as part of the Interaction Institute for Social Change (2016)'s intentional program to refocus discourse on outcomes. The Interaction Institute's

> **Commented [RM2]:** Add cultural anthro examples? Dan Sperber, and also Olivier Morin, _How Traditions Live and Die_

---

[1] See e.g. https://interactioninstitute.org/illustrating-equality-vs-equity/

version is widely distributed in educational materials teaching that 'equality' refers to opportunities while 'equity' refers to fair outcomes (Daisy, 2019), and this distinction has now made it into dictionaries used by reporters ("Equity," 2016).

Commented [RM3]: Add also Kamala's tweet.

The 'equity' episode is a clear lexical negotiation success story, even if as-yet incomplete, and it neatly straddles the various types of metalinguistic moves Hansen catalogs. It is clearly semantic/conceptual, since many of the resources explicitly claim and disclaim certain semantics for the terms 'equity' and 'equality.' The original conjunction of those terms with the boxes-and-baseball-game graphic, however, seems to have been motivated by a mere concern for snappy *usage*, within existing semantic valences. Furthermore, while many of the proposals for this distinction are explicit, most people have likely seen the image without any extensive explanation, and so accept the distinction implicitly. The initial image was designed for the Google+ community, but later versions clearly intended a society-wide change. Finally, the disputes have concerned not only which term to use, but also what images (signs, even if non-lexical) best represent the conceptual distinction. The lexical negotiation framework must be broad enough to handle all of these variations.

2. The Lexical Negotiation Model

I propose that taking such episodes seriously as *negotiations* is key to understanding when and how they succeed. Insofar as negotiations involve both an offer and its acceptance, the account is fundamentally externalist—semantics and other consequences of lexical signs are not merely a matter of the speaker's intention. The central feature of the account is that lexical negotiation is *costly* but that the cost is *variable*. When the costs are low, negotiations are likely to succeed, but when the costs are high, negotiations are likely to fail. Understanding why costs are often high shows why linguistic change is generally slow and semantic externalism feels like a binding constraint, while understanding why costs are sometimes low shows why conceptual engineering programs sometimes succeed.

The formal model has two elements. First, the linguistic consequences (e.g. semantic values) of lexical items are understood as allocations of resources. These resources may not be fully rivalrous, but they are at least partially so. Homonymy and polysemy occur, but a language with one lexical item serving for all of its words (e.g. Buffalo buffalo Buffalo…) or one word serving for all of its meanings would hardly be workable. This is clear in that many lexical negotiations, like the 'equity' and 'equality' case, explicitly work to exclude as well as include certain linguistic consequences. That allocation of linguistic resources is Pareto-efficient when it is impossible to alter those consequences to better meet one conversation partner's preferences without harming the preferences of another conversation partner. In a Pareto-efficient allocation, any change to the lexical consequences makes at least one conversation partner worse off, giving that person a reason to reject the offer.

Second, conversation partners have total control of the consequences of their lexical signs (eg, meanings) when they achieve Pareto-efficiency, since control is collective as Koch pointed out, and a Pareto-efficient allocation maximizes the achievement of conversation partners' collective desires. While Cappelen (2018) claims that semantics supervening on the partners' preferences is inadequate for control of meaning because they do not control the determination relation, as many others have noted this proves too much. Again taking the negotiation framework seriously, stock traders generally have control over the prices at which they buy and sell by placing orders to their brokers, even though they do not control or understand the deeply obscure mechanisms by which such successful negotiations are executed (Putniņš, 2010). The recent difficulties that Robinhood users had selling Gamestop stock are the exception that proves the rule (Denier, 2021). Control may not always be absolute, but if lexical negotiations were as successful as stock-trading there would be few pessimists regarding conceptual engineering, even though the determination relations are not themselves under the negotiator's control.

> **Commented [RM4]:** Who and how? Psych lit?

5

The work in translating these formal elements of the model into understanding of how difficult control of lexical consequences will be in a particular case is done by the Coase Theorem. Ronald Coase (1960) showed that the initial allocation of resources can only prevent negotiations from achieving Pareto-efficiency when transaction costs are present. If negotiation is costless for the parties, they will always achieve an agreement which is Pareto-efficient, because otherwise further negotiations would lead to an outcome where at least one party's preferences were satisfied to a greater degree and no other party's preferences fared worse than they had before the further round of negotiations. Thus in this model of lexical negotiation, the assignment of lexical consequences will be completely under the control of the conversation partners when their negotiation is free of transaction costs, but escape their control to the degree that their negotiation is itself costly. These so-called transaction costs can arise either in the performance of the negotiation itself or in the transfer of resources in accordance with the negotiated allocation. Both of these sources of transaction costs for economic negotiations have the same effect on efficiency of allocation (Frech, 1979).

There are a number of reasons to assume that lexical negotiations will generally face significant transaction costs. First, as Hansen (forthcoming) documents, proposals are generally accepted quickly—but not always. Time and attention are serious constraints on most conversations, whether everyday oral ones or those conducted through philosophy journals. Spending those resources on lexical negotiations therefore serves to distract from other communicative goals, as any philosopher who has wrangled with a reviewer over a piece of jargon can attest. Lexical negotiations therefore have serious opportunity costs. Second, new semantic associations need to be stored in a part of working memory called the phonological loop (Baddeley, 2003), which typically only holds about 26 items (Baddeley, 1992), though this is dependent on whether the new semantic associations are related to older ones (Papagno et al., 1991). Thus there may be a hard upper limit on how many lexical items in a conversation can have their consequences modulated beyond the typical range. Again, any philosopher

Commented [RM5]: Add the Relevance Theory literature.

who has tried to read a jargon-filled article from another subfield, even if all the terms are clearly defined, has experienced this pain. The typical range will vary from conversation partner to conversation partner, but it is a binding constraint beyond the scope of any particular interaction, and thus serves as a transaction cost on lexical negotiations. Third, sometimes negotiators can fail to realize that their offers have not been accepted, leading to conversational breakdowns. While this 'transformative communicative disruption' may actually be a benefit for serious conceptual engineering projects due to the personal transformation it induces (Sterken, 2019), it also obviously limits the number of such projects one may wish to concurrently undertake.

Other transaction costs have a less cognitive character. A fourth cause of (especially) costly negotiations is what Cappelen (2018) calls 'lexical effects'—'the non-cognitive, non-semantic, non-pragmatic effects of words.' These can potentially be negotiated away in the long term—slurs do sometimes lose their power—but in the immediate case they serve as a serious cost of the negotiation itself. As the recent case of New York Times reporter Donald McNeil illustrates, these effects can actually be direct results of the lexical sign used (Kelly & Italiano, 2021) and thereby escape the use/mention distinction. Negotiations typically involve actual use of a term (Hansen, forthcoming), but this means that even mention will bear significant costs. Techniques for negotiation without mention have developed, like grawlix (Merriam Webster, 2018),[2] but those increase the cognitive costs—as you are probably experiencing right now with 'grawlix.' These lexical effects are compounded by their uncertainty—most people react poorly to slurs, but the degree of negative reaction varies quite widely. Volatility adversely affects prices, so negotiations with uncertain transaction costs are markedly less efficient than the absolute size of those costs would otherwise indicate (Ho et al., 1985).

---

[2] I.e., use of asterisks and other punctuation to represent forbidden lexical signs.

A fifth source of transaction costs for lexical negotiations is that the existing four can be self-compounding, like a stock bubble. Extensive conversational moves which tax our conversational resources often read as obscure or prolix. These features, however, are ones which H.P. Grice (1967) identified as indicating conversational implicature. Rather than perceiving costly negotiations as being in good faith, conversation partners may instead read such offers as indicating a desire to condescend, change the subject, or otherwise offend. Anyone who has witnessed an internet comment thread between two interlocutors with wildly different baseline meanings can attest to this effect, which usually results in negotiation being halted entirely rather than succeeding at all, let alone with Pareto-efficiency.

The net result of all of these sources of transaction costs is that lexical negotiations rarely achieve Pareto-efficient results. There are often further moves left on the table which would benefit one or both conversation partners without cost to the other, but the negotiation costs themselves prevent further searching for such solutions. Philosophical debates over certain terms are hot for a while, but then fade while another set of terms in the same neighborhood heats up, as in the move from 'reduction' to 'supervenience' to 'grounding' talk.[3] Economic research has shown transaction costs are negatively associated with the number of changes in allocation and positively associated with the stability of allocations (Barclay et al., 1998). Big shifts in lexical consequences might have big rewards, but they also come with correspondingly high transaction costs, so changes will be smaller and more infrequent than negotiators might wish.

These general considerations help to explain the success of the 'equity' example. The offer of lexical allocation was made with a high production quality professionally drawn cartoon, minimizing the

---

[3] I don't mean that these terms are identical in meaning, or that the disputes surrounding them are merely verbal, but rather that the decision to switch among these terms rather than further qualifying the prior ones which have overlapping meanings (Gillett & Aizawa, 2016) is a decision to close one set of negotiations and open another.

8

cognitive costs needed for processing. The negotiation involved only two terms, and while the distinction between them was novel, neither was moved far outside of its previous semantic range. 'Equity' is a familiar enough term that it is not difficult to remember or manipulate, but sufficiently outside of everyday vocabulary that its existing resonances are not too strong. Both 'equality' and 'equity' have positive valences, so no negative lexical effects were in play. The Interaction Institute for Social Change backed the image with institutional support to spread educational materials, so that communicative confusions would be addressed in a systematic and ongoing way. At the same time, the image primarily spread through social media, so its negotiation was pitched within existing communities of interlocutors, minimizing any misleading conversational implicature. Transaction costs likely remain present, but the conditions for success were present.

Conclusions

Generally high transaction costs in a lexical negotiation framework validate the features of language frequently remarked on by philosophers. Modest semantic externalism obtains in this framework because the meaning of words used in a conversation is partly dependent on factors beyond the control of any of the conversation partners—the pre-negotiation semantics. Orthography, paleography, and pronunciation will be largely stable for similar reasons. This explains the general stability of language across time and context remarked on by Cappelen. Successful negotiation requires being perceived as an authority by one's interlocutors (Lewis, 2020), so marginalized persons will be at a disadvantage in metalinguistic discourse. Tyler Burge (1979)'s ill-educated patient loses the negotiation over the meaning of arthritis not only because his view is idiosyncratic but because he is not a doctor. Arthritis, after all, is as Burge says not a natural kind.

> **Commented [RM6]:** Summarize the argument and situate it w/r/t Maitra (2012). What are the limits to this?

In tightly-knit groups, by contrast, reactions will be more predictable, good-will more assured, and negotiations more iterative, so transaction costs should be low. This means that lexical negotiations

within such groups should more nearly approach Pareto-efficiency, and larger swings in semantic allocations should be observed. This is empirically compatible with modest transaction costs: research shows that intra-household monetary allocations are approximately Pareto-efficient (Browning & Chiappori, 1998), even though anyone who has negotiated with a domestic partner knows that doing so is not cost-free. Such small, intimate groups may thus frequently create micro-languages with significant semantic drift. Indeed linguists call such subgroup-specific micro-languages 'registers' and one of the factors that distinguishes them from sociolects is their semantic modulation (Lewandowski, 2010). Anyone who has tried to make sense of a teenager's text messages can attest to the difficulty of interpreting such communication as an outsider. This substantiates Peter Ludlow (2014)'s view of the pervasiveness of such small-scale control without suggesting that it should be equally effective at larger scales, since negotiation costs scale super-linearly with network size (Scholtes et al., 2016). Such lexical negotiations are thus a possible basis for handling indexicals and the Paderewski puzzle as Ludlow suggests.

This framework also has implications for evaluating the likelihood of success in conceptual engineering projects. Kevin Scharp (2021)'s work on 'truth' is promising because it is largely addressed to a small subfield, easing the cognitive constraints, and truth paradoxes are enduring subjects of philosophical reflection. Scharp also does a great deal to persuade his readers that existing solutions are not already Pareto-efficient, since each view of truth leads to frustrated desires. Haslanger's program seems likely to face greater challenges. 'Woman' is already an extremely well-established lexical sign among a large population with strong emotional resonances. Further, the term 'subordination' which Haslanger wishes to introduce into the definition of 'woman' has strong negative resonances, exacerbated by the unpredictability of reactions to discussions of gender. Finally, the lexical consequences of 'woman' may already be Pareto-efficient since there are large and well-organized institutions opposed to any change

10

in its semantic value. The implication is not that philosophers with such difficult projects should give up, but rather that they will require a great deal of patience and social solidarity.

3. References

Baddeley, A. (1992). Working Memory. *Science*, *255*(5044), 556–559. JSTOR.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Barclay, M. J., Kandel, E., & Marx, L. M. (1998). The Effects of Transaction Costs on Stock Prices and Trading Volume. *Journal of Financial Intermediation*, *7*(2), 130–150. https://doi.org/10.1006/jfin.1998.0238

Browning, M., & Chiappori, P. A. (1998). Efficient Intra-Household Allocations: A General Characterization and Empirical Tests. *Econometrica*, *66*(6), 1241–1278. JSTOR. https://doi.org/10.2307/2999616

Burge, T. (1979). Individualism and the Mental. *Midwest Studies In Philosophy*, *4*(1), 73–121. https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

Cappelen, H. (2018). *Fixing Language. An Essay on Conceptual Engineering*. Oxford Univ. Press. http://www.oxfordscholarship.com/view/10.1093/oso/9780198814719.001.0001/oso-9780198814719

Chalmers, D. J. (2011). Verbal Disputes. *The Philosophical Review*, *120*(4), 515–566.

Coase, R. H. (1960). The Problem of Social Cost. *The Journal of Law and Economics*, *3*, 1–44. https://doi.org/10.1086/466560

Daisy. (2019, March 29). Equality and Equity. *Useful Stuff for Changemakers*. https://social-change.co.uk/blog/2019-03-29-equality-and-equity

Denier, A. (2021, January 28). *Webull CEO: Here's why Robinhood is restricting users from buying stocks like GameStop* (Yahoo! Finance, Interviewer) [Live Video].

https://finance.yahoo.com/video/heres-why-robinhood-restricting-users-173049721.html

Equity. (2016). In *The Glossary of Education Reform*. Great Schools Partnership.

https://www.edglossary.org/equity/

Espinoza, O. (2007). Solving the equity–equality conceptual dilemma: A new model for analysis of the educational process. *Educational Research*, *49*(4), 343–363.

https://doi.org/10.1080/00131880701717198

Frech, H. E. (1979). The Extended Coase Theorem and Long Run Equilibrium: The Nonequivalence of Liability Rules and Property Rights. *Economic Inquiry*, *17*(2), 254–268.

Fricker, M. (2006). Powerlessness and Social Interpretation. *Episteme*, *3*(1–2), 96–108.

https://doi.org/10.3366/epi.2006.3.1-2.96

Froehle, C. (2020, June 7). The Evolution of an Accidental Meme [Medium]. *CRAIG*.

https://medium.com/@CRA1G/the-evolution-of-an-accidental-meme-ddc4e139e0e4

Gillett, C., & Aizawa, K. (Eds.). (2016). *Scientific composition and metaphysical ground* (1st edition). Springer Berlin Heidelberg.

Grice, H. P. (1967). Logic and Conversation. In *Studies in the Way of Words* (pp. 41–58). Harvard University Press.

Hansen, N. (forthcoming). Metalinguistic Proposals. *Inquiry*.

Haslanger, S. (2000). Gender and Race: (What) Are They? (What) Do We Want Them to Be? *Noûs*, *34*(1), 31–55.

Ho, T. S. Y., Schwartz, R. A., & Whitcomb, D. K. (1985). The Trading Decision and Market Clearing Under Transaction Price Uncertainty. *The Journal of Finance*, *40*(1), 21–42. JSTOR.

https://doi.org/10.2307/2328046

Illustrating Equality VS Equity. (2016, January 14). Interaction Institute for Social Change.

    https://interactioninstitute.org/illustrating-equality-vs-equity/

Kelly, K. J., & Italiano, L. (2021, February 5). Two New York Times staffers resign, citing inappropriate

    behavior. *New York Post*. https://nypost.com/2021/02/05/new-york-times-staffers-donald-

    mcneil-andy-mills-resign/

Koch, S. (2021). The externalist challenge to conceptual engineering. *Synthese*, *198*(1), 327–348.

    https://doi.org/10.1007/s11229-018-02007-6

Lewandowski, M. (2010). Sociolects and Registers – a Contrastive Analysis of Two Kinds of Linguistic

    Variation. *Investigationes Linguisticae*, *20*, 60. https://doi.org/10.14746/il.2010.20.6

Lewis, K. S. (2020). The Speaker Authority Problem for Context-Sensitivity (Or: You Can't Always Mean

    What You Want). *Erkenntnis*, *85*(6), 1527–1555. https://doi.org/10.1007/s10670-018-0089-2

Ludlow, P. (2014). *Living Words: Meaning Underdetermination and the Dynamic Lexicon*. Oxford

    University Press. https://doi.org/10.1093/acprof:oso/9780198712053.001.0001

Merriam Webster. (2018, June). What the @% Is a 'Grawlix'? *Words We're Watching*.

    https://www.merriam-webster.com/words-at-play/grawlix-symbols-swearing-comic-strips

Miller, J. T. M. (2020). On the individuation of words. *Inquiry*, *63*(8), 875–884.

    https://doi.org/10.1080/0020174X.2018.1562378

O'Conner, P. T., & Kellerman, S. (2019, June 14). The Grammarphobia Blog: On 'equity' and 'equality.'

    Grammarphobia. https://www.grammarphobia.com/blog/2019/06/equity-equality.html

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language

    vocabulary learning. *Journal of Memory and Language*, *30*(3), 331–347.

    https://doi.org/10.1016/0749-596X(91)90040-Q

Plunkett, D. (2015). Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology

    of Philosophy. *Inquiry*, *58*(7–8), 828–874. https://doi.org/10.1080/0020174X.2015.1080184

Plunkett, D., & Sundell, T. (2013). Disagreement and the Semantics of Normative and Evaluative Terms. *Philosopher's Imprint*, *13*(23). http://hdl.handle.net/2027/spo.3521354.0013.023

Putniņš, T. (2010). Naked short sales and fails-to-deliver: An overview of clearing and settlement procedures for stock trades in the USA. *Journal of Securities Operations & Custody*, *2*(4), 340–350.

Scharp, K. (2021). Conceptual engineering for truth: Aletheic properties and new aletheic concepts. *Synthese*, *198*(2), 647–688. https://doi.org/10.1007/s11229-019-02491-4

Scholtes, I., Mavrodiev, P., & Schweitzer, F. (2016). From Aristotle to Ringelmann: A large-scale analysis of team productivity and coordination in Open Source Software projects. *Empirical Software Engineering*, *21*(2), 642–683. https://doi.org/10.1007/s10664-015-9406-4

Sterken, R. K. (2019). Linguistic Interventions and Transformative Communicative Disruption. In H. Cappelen, D. Plunkett, & A. Burgess (Eds.), *Conceptual Engineering and Conceptual Ethics*. Oxford: Oxford University Press.