Journal of Memory and Language 78 (2015) 1–17

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Talkers account for listener and channel characteristics to communicate efficiently

John K Pate [a,*], Sharon Goldwater [b]

[a] Department of Computing, Macquarie University, Sydney, NSW 2109, Australia[1]
[b] School of Informatics, The University of Edinburgh, 10 Crichton St., Edinburgh EH8 9AB, UK

ARTICLE INFO

Article history:
Received 2 October 2013
revision received 10 October 2014
Available online 29 October 2014

Keywords:
Language production
Predictability effects
Language production efficiency
Uniform information density
The smooth signal hypothesis

ABSTRACT

A well-known effect in speech production is that more predictable linguistic constructions tend to be reduced. Recent work has interpreted this effect in an information-theoretic framework, proposing that such *predictability effects* reflect a tendency towards communicative efficiency. However, others have argued that these effects are, in the terminology of Gould and Lewontin (1979), *spandrels*: incidental by-products of other processes (such as a talker-oriented tendency for low production effort). This article develops the information-theoretic framing more fully, showing that information-theoretic efficiency involves different kinds of coding operations (predictability effects), not all of which are consistent with the spandrel account. Using mixed effects regressions, we analyze word durations in several spontaneous speech corpora, comparing predictability effects between infant-directed and adult-directed speech and between speech to visible and invisible listeners. We find that talkers adjust the extent to which production varies with predictability measures according to listener characteristics, and exploit an additional visual channel to eliminate phonetic redundancy. This pattern would demand multiple independent spandrel accounts, but is unified by an adaptive account. Our results broaden the scope of existing work on predictability effects and provide further evidence that these effects are tied to communicative efficiency.

Introduction

Talkers can usually communicate a given message using many different linguistic forms: for a fixed meaning or intention, talkers can produce different syntactic constructions, select different words with similar meanings, and pronounce parts of the utterance more or less clearly. Some of these productions will be longer and more distinct (overarticulated), while others will be shorter and more ambiguous (reduced). For example, the pronunciation of *feline* typically takes longer than a pronunciation of *cat*, but the pronunciation of *cat* is more ambiguous (e.g. its phonological form is a substring of many other words).

Many researchers have argued that talkers make these choices, whether consciously or unconsciously, in ways that lead to more efficient communication. Lindblom (1990), focusing on word production, observed that speech appeared to vary on virtually every acoustic dimension according to a range of variables, including local predictability, talker identity, listener needs, and so forth. Lindblom proposed that speech lacked reliable invariants because the goal of speech is not the approximation of some ideal, but discrimination among different items in a lexicon. His **H**yper- & **H**ypo-articulation (H&H) theory proposed that talkers articulate "just enough," in terms of

[1] Present address.
* Corresponding author.
E-mail addresses: john.pate@mq.edu.au (J.K Pate), sgwater@inf.ed.ac.uk (S. Goldwater).

http://dx.doi.org/10.1016/j.jml.2014.10.003
0749-596X/

effort or acoustic distinctiveness, to enable reliable discrimination. Words that can be discriminated from other possible words on the basis of non-acoustic grounds require less acoustic distinctiveness, and so less articulatory effort, while talkers must provide more acoustic information for words that are harder to identify.

Subsequent proposals (e.g. Aylett & Turk, 2004; Jaeger, 2010) have sought to elaborate what "just enough" means in different ways, which we will discuss shortly, but they all similarly hinge on the observation that talkers tend to produce more reduced (more ambiguous) forms for portions of their message that are more probable. Such *predictability effects* are pervasive throughout language, resulting in shorter phonological forms for more common words (Zipf, 1949), phonetic reduction of words that are discourse given or have higher *n*-gram probabilities (Aylett & Turk, 2004; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Seyfarth, 2014; Priva, 2008), and reduction or omission of words that have higher syntactic probabilities (Frank & Jaeger, 2008; Gahl & Garnsey, 2004; Gahl, Garnsey, Fisher, & Matzen, 2006; Levy & Jaeger, 2007; Tily et al., 2009). Broadly, predictability effects are information-theoretically efficient because they make shorter utterances more common than longer ones while guarding against communication errors.

Although previous work has shown that predictability effects exist, and that they make speech more efficient, it has not been established that predictability effects are an adaptation for efficient communication for two reasons. First, it is difficult to measure *how much* more efficient communication is as a result of predictability effects, and so we do not know if they have any practical impact. Second, it is possible to derive at least some predictability effects as a consequence of how lexical and grammatical knowledge might be stored and accessed. For example, Dell and Brown (1991) and Ferreira (2003) have highlighted the possibility that some aspects of linguistic knowledge and processing are shared between production and comprehension. With this kind of sharing, the same items that are hard to access in comprehension will be hard to access for production. If there is also some mechanism that lengthens the realization of hard-to-access items, talkers would end up producing longer realizations for items that would be hard for them to comprehend, if they were the listener. However, this situation would not necessarily lead to a practical improvement in communication rate or reliability: the magnitude of this effect is not tuned for efficiency but is simply a consequence of whatever system lengthens the realization of hard-to-access items. Thus, if predictability effects make communication only negligibly more efficient, and result from architectural accidents, they cannot be said to reflect adaptation to communicative efficiency in any real sense.

This example unearths the worry that predictability effects may be what Gould and Lewontin (1979) term "evolutionary spandrels," by analogy with a particular architectural feature of buildings with arches. Gould and Lewontin observed that the most striking features of arches are the designs and mosaics that appear in the roughly triangular region, called the spandrel, between arches at right angles. Despite the striking appearance, the spandrels are not a driving feature of the architecture: they exist because arches must be curved to support weight. Gould and Lewontin (1979) used this example to illustrate the point that even very striking biological features may be evolutionary coincidences or by-products, without being adaptive themselves. A general correlation between distinctiveness, or ease of perception, and coarse measures of predictability is not sufficient to establish that language is adapted for communicative efficiency, no matter how tantalizing the prospect.

One especially plausible spandrel account appeals to production or planning difficulty: people avoid productions that require more effort, and these same productions tend to be rare. Under this account, production difficulty alone is the driving force, and there should not be an independent effect of predictability. However, recent experimental work has found an independent contribution of predictability, after controlling for production difficulty (Jaeger, 2013). For example, Kurumada and Jaeger (2013) examined the production of optional case markers in Japanese that are sometimes redundant with other grammatical function cues, and found independent contributions of both production difficulty and predictability. Baese-Berk and Goldrick (2009), with replications using different methodology by Peramunage, Blumstein, Myers, Goldrick, and Baese-Berk (2011), Kirov and Wilson (2012), Buz, Jaeger, and Tanenhaus (2014), looked at voice-onset time (VOT), an important cue to voicing contrasts in English stops. They found that talkers produce a stronger VOT contrast when both words of a voicing minimal pair have plausible referents in the current discourse context. Buz et al. (2014) additionally looked for an effect of planning difficulty (as measured by production latency), but did not find one. Although this last result, as all null results, should be interpreted with caution, these studies together begin to provide evidence that predictability effects on at least morphology and phonetic detail cannot be solely attributed to production difficulty.

This paper provides a complementary line of evidence favoring the view that predictability effects reflect an adaptation towards communication. We investigate how distinctiveness correlates with different sources of predictability. These different sources are unified under an account that appeals to information-theoretic efficiency, but, by virtue of their differences, would presumably require independent spandrel accounts. To do this, we first develop the information-theoretic framework of recent work on predictability effects (Aylett & Turk, 2004; Jaeger, 2010; Levy & Jaeger, 2007) to classify predictability effects into three kinds: source coding from the talker's perspective, source and channel coding specialized to listener characteristics, and channel coding. These three kinds of predictability effects consider very different features, and so would presumably involve independent spandrel accounts. They are unified, however, by their relevance to information-theoretic optimization. With this theoretical framework in mind, we provide two corpus studies that find evidence for the latter two kinds of predictability effects. Concretely, we use mixed-effects regression models to analyze several corpora of spontaneous speech, examining predictability measures that have been