



What the Heck Is Saliency? How Predictive Language Processing Contributes to Sociolinguistic Perception

T. Florian Jaeger^{1,2,3*} and Kodi Weatherholtz¹

¹ Human Language Processing Lab, Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA, ² Department of Computer Science, University of Rochester, Rochester, NY, USA, ³ Department of Linguistics, University of Rochester, Rochester, NY, USA

Keywords: accent, dialect, idiolect, saliency, surprisal, prediction, expectation, learning

OPEN ACCESS

Edited by:

Adriana Hanulíková,
University of Freiburg, Germany

Reviewed by:

Lauren Squires,
Ohio State University, USA
Christian Langstrof,
University of Freiburg, Germany
Daniel Müller-Feldmeth,
University of Freiburg, Germany

*Correspondence:

T. Florian Jaeger
tjaeger@ur.rochester.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

Received: 07 April 2016

Accepted: 12 July 2016

Published: 03 August 2016

Citation:

Jaeger TF and Weatherholtz K (2016)
What the Heck Is Saliency? How
Predictive Language Processing
Contributes to Sociolinguistic
Perception. *Front. Psychol.* 7:1115.
doi: 10.3389/fpsyg.2016.01115

INTRODUCTION

Some sociolinguistic variables are prone to hypercorrection, stigmatization and style shifting, while other variables are not. The status of the former type—sometimes called *stereotypes* and *markers* (Labov, 1972)—has been attributed to the increased meta-linguistic awareness language users seem to have of these variables. This awareness in turn is attributed to the *saliency* of these variables, such that greater saliency is assumed to cause greater meta-linguistic awareness (e.g., Trudgill, 1986). Saliency has similarly been invoked when aiming to explain implicit social inferences about, or attitudes toward, speakers who exhibit certain variables in their speech (Babel, 2016; Drager and Kirtley, 2016; Squires, 2016). However, saliency is a hard to define concept (for review, see Auer et al., 1998; Kerswill and Williams, 2002) and, partly as a consequence, “notoriously difficult to quantify” (Hickey, 2000). For a concept that plays such a central and ubiquitous role in sociolinguistic explanations, this is arguably a dangerous state of affairs.

This motivates the present commentary. We believe that advances in computational psycholinguistics offer definitions of sociolinguistic saliency that are more concrete, both empirically and formally grounded, and quantifiable (and thus falsifiable). We propose that it is important to distinguish between the *initial saliency* a listener experiences when first encountering a novel variant (e.g., because of exposure to a previously unfamiliar dialect, sociolect, or idiolect—henceforth *lects*; Schirmunski, 1930; Preston, 1996), and saliency at later stages. Saliency after the initial encounter is the cumulative product of an individual’s experience related to the lectal variant, including direct experience, as well as discourse about the variant (e.g., explicit stereotyping or enregisterment, Agha, 2003). Here we focus on the causes for initial saliency, which we think *can* be defined in a principled and quantifiable way.

Specifically, we propose that saliency in the first moments when a novel lect is encountered cannot be understood without reference to prior expectations based on listeners’ past language experience and the ensuing expectation violation that a listener experiences relative to those prior expectations—an idea explored in more depth by Rác (2012, 2013). Here we contribute to these efforts. We draw on basic concepts from probability and information theory to define initial saliency as a function of (top-down) prior expectations. This has several advantages. First, the proposed definition of saliency is quantifiable (see also Rác, 2013). Second, computational psycholinguistics has linked the very same quantities to language processing and learning. Recognizing this link offers the opportunity to ground sociolinguistic saliency in human information processing—both empirically and theoretically—offering a parsimonious account of initial saliency.

After we have outlined our proposal, we briefly turn to an apparent puzzle that was raised during the workshop leading to this special issue: several presenters pointed out that saliency sometimes seems to be inversely related to the frequency of a variant and other times positively related. This puzzle readily dissolves once the view proposed here is taken into account.

FIRST ENCOUNTERS WITH A VARIANT: SURPRISAL AS A MEASURE OF INITIAL SALIENCY

Imagine a listener during the first moments of encountering a talker who speaks in an unfamiliar lect. The unfamiliar lect by definition differs from what the listener has previously experienced. Following the sociolinguistic literature, we can think of these differences as differences in the realization of linguistic variables, and the specific realization of the variables as lectal variants (Labov, 1966). What then makes a lectal variant salient in this hypothetical first encounter? Research in sociolinguistics has identified a number of perceptual features that can contribute to the perception of a variant as salient, such as a priori perceptual or articulatory distinctiveness (for review, see Auer et al., 1998). However, influences of prior experience are arguably as important or more important. Specifically, variants that are *unexpected* given the listener's *prior expectations about linguistic variables* (including, broadly speaking, the listener's language background) should be more salient in the moment they are experienced.

Events that we do not expect, or that are surprising to us, tend to stand out. There is now strong evidence that this anecdotal observation about strongly unexpected events extends to subtle and highly gradient differences in unexpectedness. During language processing, words and structures that are less expected are processed more slowly (e.g., MacDonald et al., 1994; Garnsey et al., 1997; McRae et al., 1998; MacDonald and Shillcock, 2003) and they are recognized accurately less often in noise (Cole and Perfetti, 1980; Grosjean, 1980). Critically, similar costs of unexpectedness are observed for unfamiliar lectal variants when comprehenders first encounter them (e.g., Kaschak and Glenberg, 2004; Squires, 2014a; Fraundorf and Jaeger, in press). Unexpectedness—or the degree to which something is violating our expectations based on previous experience—can be measured in a number of ways. One principled measure is referred to as *surprisal* (Hale, 2001; Levy, 2008). The surprisal associated with processing a certain input (e.g., a phonetic feature, phonological category, word, or syntactic structure) is identical to the amount of new information gained by processing the input, also known as the Shannon information (Shannon, 1948).

The surprisal of a unit is defined as the logarithm of the inverse of the contextual probability of the unit:

$$I(\text{unit}) = \log \frac{1}{p(\text{unit} | \text{context})} = -\log p(\text{unit} | \text{context}) \quad (1)$$

If the logarithm of the inverse contextual probability is taken to base 2, surprisal measures the number of bits of information

gained by processing the input over and above what was expected prior to processing the input. The surprisal of a word in (linguistic) context has been found to be proportional to its average reading time (Frank and Bod, 2011; Smith and Levy, 2013; Linzen and Jaeger, 2016). Surprisal has also been found to be correlated with neural signatures in ERP or MEG studies (Frank et al., 2015; for further references, see Kuperberg and Jaeger, 2016).

Recent studies have further linked surprisal to implicit learning operating during language processing (Fine and Jaeger, 2013; Jaeger and Snider, 2013; for a related view, see Dell and Chang, 2014). As is well-known from sociolinguistic research, talkers differ in their pronunciation, lexical, and syntactic preferences (among other things, Labov, 1972). As a consequence, efficient and robust language processing requires that linguistic expectations need to flexibly adapt to these differences (Fine et al., 2013; Kleinschmidt and Jaeger, 2015). Indeed, expectation adaptation has now been documented for speech perception (Clayards et al., 2008; for review, see Weatherholtz and Jaeger, in press), lexical (Creel et al., 2008), syntactic (Fine et al., 2013), and prosodic processing (Kurumada et al., under review), including adaptation to novel lectal variants (e.g., Kaschak and Glenberg, 2004; Bradlow and Bent, 2008; Kraljic et al., 2008; Fraundorf and Jaeger, in press). Adaptation to changes in the statistics of the environment should be sensitive to surprisal (or more generally to expectation violation): the degree to which inputs differ from prior expectations is informative about how and how much learners need to adapt their future expectations (Courville et al., 2006; Qian et al., 2012). Consistent with this prediction, there is evidence that the amount of expectation adaptation after processing unexpected linguistic input is proportional to that input's surprisal (Fine and Jaeger, 2013; Arai and Mazuka, 2014; for related evidence from production, see Berolet and Hartsuiker, 2010; Jaeger and Snider, 2013).

Taken together, this research suggests that surprisal (or its generalization, Bayesian surprise; Itti and Baldi, 2009) is a plausible measure of “unexpectedness” and, as such, one factor that is likely to contribute to the initial saliency of newly encountered lectal variants. Specifically, it is the surprisal of the variant *given the prior expectations of the listener* that is expected to predict initial saliency. These prior expectations, we further submit, depend not only on *linguistic context* (e.g., the probability of a lectal variant given surrounding phonological or lexical information, including the presence or absence of other lectal variants) but also on *social context* (e.g., the probability of a lectal variant given socio-indexical information about the talker).

Consider, for example, a specific linguistic variable, such as /t/-deletion or flapping: if this variant occurs overall much more frequently in a newly encountered lect than a priori expected or in different phonological and lexical contexts than a priori expected, it will have high surprisal (this reasoning also extends to novel, not previously encountered, variables)¹. It is in this

¹Under the naïve assumption that everything that has never been observed is considered to have a probability of 0, the surprisal of a novel variant would be infinite. This is avoided, if some probability mass is held out to account for the fact that we do, in fact, observe novel events even as adults.

sense that the saliency of a lectal variant is *inversely* related to frequency—specifically to the *expected* relative contextual frequency of the variant².

Since the expectations that determine the surprisal of a lectal variant reflect the individual's previous language experience, it naturally follows that initial saliency can be “different for different social groups” (Kerswill and Williams, 2002) and individuals (see also Hickey, 2000; Campbell-Kibler, 2012). Specifically, initial saliency should depend on which lects the individual has previously been exposed to, the frequency of the novel lectal variant in those familiar lect, and perhaps the frequency of similar variants in familiar lects (see Squires, 2014b). Next we turn to the question of how the initial saliency of a variant is related to the probability that the variant will become associated with the lect, thereby acquiring social meaning.

BEYOND THE FIRST ENCOUNTER: FREQUENCY AND ASSOCIATION

What then happens over time, as a novel lectal variant is encountered again? Consider a novel talker producing a high surprisal variant only once, compared to producing that (equally high surprisal) variant repeatedly. Intuitively, listeners should be more likely to learn an association between the variant and the novel lect in the latter case: while the surprisal of a lectal variant determines how much it “stands out,” the frequency with which the lectal variant is observed increases the probability that the variant is perceived and learned—a prerequisite to becoming associated with the lect. It is in this sense that the resulting sociolinguistic saliency of a variant is *positively* related to its (*actually observed* relative) frequency *in the novel lect*. Note that this is not in conflict with our previous statement. Surprisal is predicted to cause the initial saliency experienced when observing a lectal variant that was unexpected based on prior experience. High frequency in the novel lect—or specifically the cumulative effect of the surprisal experienced whenever a variant is encountered again—is predicted to increase the likelihood that the listener learns that the variant is associated with the lect (this idea is closely related to the *mutual information* between the variant and lect).

This also predicts that lectal variants can become associated almost instantaneously with a new lect or social group *if the variant is particularly unexpected* (as seems to be the case, Squires, 2014a). Such *ad-hoc* associations should be even more likely when listeners have other reasons to believe (rightly or wrongly) that the producer belongs to a novel group—a prediction that, to the best of our knowledge has not been directly tested.

Viewed this way, we can think of the sociolinguistic saliency that a lectal variant acquires over time as being a function of its (perceived) informativeness about social group membership.

²There is one caveat to this prediction: prior expectations also affect *what* we perceive (cf. perceptual illusions or the perceptual magnet effect; Kuhl, 1991), and therefore can lead to a non-faithful representation of the perceptual input (cf. Feldman et al., 2009).

This raises an interesting question for future research. There is now evidence that listeners develop and store implicit models or expectations about different lects that they have been exposed to (Niedzielski, 1999; Strand, 1999; Bradlow and Bent, 2008; Walker and Hay, 2011; Hanulíková et al., 2012; Shaw et al., 2015; for review, see Foulkes and Hay, 2015; Kleinschmidt and Jaeger, 2015). It is, however, still an open question to what extent the features that these implicit expectations are conditioned on are the same that more explicit processes, such as stereotyping refer to.

CONCLUSION

We propose that research on sociolinguistic saliency needs to take into account what is known about language processing and learning (see also Rácz, 2013; for a related perspective that grew out of the same workshop, see Schmid and Günther, 2016). One consequence of this is that the surprisal and frequency of lectal variants are likely predictors of a variant's saliency. Specifically, surprisal is high when first encountering unfamiliar lectal variants. With further exposure, the association between the variant and the lect increases, while the surprisal evoked by the variant decreases.

One advantage of this approach to saliency is that it makes novel testable predictions, some of which we have derived above. A second benefit is that surprisal and frequency are quantitative measures that can—in principle (provided suitable corpora)—be estimated objectively from language database. Of course, other properties of lectal variants (e.g., differences in a priori perceptual saliency, such as loudness) or processes operating over them are likely to affect saliency (e.g., enregisterment, which will selectively strengthen the associations between a lectal variant and the lect; Agha, 2003; Schmid, 2007). However, these other contributors to saliency are generally difficult to measure reliably. We thus submit that the proposal outlined here should be taken into account *first*, providing a baseline for a variant's expected saliency.

AUTHOR CONTRIBUTIONS

All authors listed, have made equally substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We thank Alice Blumenthal-Dramé, Adriana Hanulíková, and Bernd Kortmann for organizing the workshop that led to this special issue, *Perceptual linguistic saliency: Modeling causes and consequence* held in Freiburg, October 15th to 17th 2014. The ideas expressed here benefitted from stimulating discussions with participants at the workshop and from the reviewers' comments, who went beyond the expected. Work on this paper was partially supported by NSF CAREER award IIS-1150028 and NICHD grant R01 HD075797 to TFJ. The views expressed here are not necessarily those of the funding agencies.

REFERENCES

- Agha, A. (2003). The social life of cultural value. *Lang. Commun.* 23, 231–273. doi: 10.1016/s0271-5309(03)00012-0
- Arai, M., and Mazuka, R. (2014). The development of Japanese passive syntax as indexed by structural priming in comprehension. *Q. J. Exp. Psychol.* 67, 60–78. doi: 10.1080/17470218.2013.790454
- Auer, P., Birgit, B., and Beate, G. (1998). Subjective and objective parameters determining 'saliency' in long-term dialect accommodation. *J. Sociolinguist.* 2, 163–187. doi: 10.1111/1467-9481.00039
- Babel, A. (ed.). (2016). *Awareness and Control in Sociolinguistic Research*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781139680448
- Bernolet, S., and Hartsuiker, R. (2010). Does verb bias modulate syntactic priming? *Cognition* 114, 455–461. doi: 10.1016/j.cognition.2009.11.005
- Bradlow, A. R., and Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition* 106, 707–729. doi: 10.1016/j.cognition.2007.04.005
- Campbell-Kibler, K. (2012). Contestation and enregisterment in Ohio's imagined dialects. *J. Eng. Linguist.* 40, 281–305. doi: 10.1177/0075424211427911
- Clayards, M., Tanenhaus, M. K., Aslin, R. N., and Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition* 108, 804–809. doi: 10.1016/j.cognition.2008.04.004
- Cole, R. A., and Perfetti, C. A. (1980). Listening for mispronunciations in a children's story: the use of context by children and adults. *J. Verb. Learn. Verb. Behav.* 19, 297–315. doi: 10.1016/S0022-5371(80)90239-X
- Courville, A., Nathaniel, D. D., and David, S. T. (2006). Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* 10, 294–200. doi: 10.1016/j.tics.2006.05.004
- Creel, S. C., Aslin, R. N., and Tanenhaus, M. K. (2008). Heeding the voice of experience: the role of talker variation in lexical access. *Cognition* 106, 633–664. doi: 10.1016/j.cognition.2007.03.013
- Dell, G. S., and Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20120394. doi: 10.1098/rstb.2012.0394
- Drager, K., and Kirtley, J. (2016). "Awareness, saliency, and stereotypes in exemplar-based models of speech production and perception," in *Awareness and Control in Sociolinguistic Research*. ed A. Babel (Cambridge, UK: Cambridge University Press), 1–24. doi: 10.1017/CBO9781139680448.003
- Feldman, N. H., Griffiths, T. L., and Morgan, J. L. (2009). The influence of categories on perception: explaining the perceptual magnet effect as optimal statistical inference. *Psychol. Rev.* 116, 752–782. doi: 10.1037/a0017196
- Fine, A., and Jaeger, T. F. (2013). Evidence for implicit learning in syntactic comprehension. *Cogn. Sci.* 37, 578–591. doi: 10.1111/cogs.12022
- Fine, A. B., Florian Jaeger, T., Thomas, A. F., and Ting, Q. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE* 8:e77661. doi: 10.1371/journal.pone.0077661
- Foulkes, P., and Hay, J. (2015). "The emergence of sociophonetic structure," in *The Handbook of Language Emergence*, eds B. MacWhinney and W. O'Grady (Hoboken, NJ: John Wiley and Sons, Inc.), 292–313.
- Frank, S. L., Leun, J. O., Giulia, G., and Gabriella, V. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain Lang.* 140, 1–11. doi: 10.1016/j.bandl.2014.10.006
- Frank, S. L., and Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychol. Sci.* 22, 829–834. doi: 10.1177/0956797611409589
- Fraundorf, S., and Jaeger, T. F. (in press). Readers generalize adaptation to newly-encountered dialectal structures to other unfamiliar structures. *J. Mem. Lang.* doi: 10.1016/j.jml.2016.05.006
- Garnsey, S. M., Neal, J. P., Elizabeth, M., and Melanie, L. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *J. Mem. Lang.* 37/1, 58–93. doi: 10.1006/jmla.1997.2512
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Percept. Psychophys.* 28, 267–283. doi: 10.3758/BF03204386
- Hale, J. (2001). "A probabilistic Earley parser as a psycholinguistic model," in *Proceedings of NAACL* (Pittsburgh, PA). doi: 10.3115/1073336.1073357
- Hanulíková, A., van Alphen, P. M., van Gochnd, M. M., and Andrea, W. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *J. Cogn. Neurosci.* 24, 878–887. doi: 10.1162/jocn_a_00103
- Hickey, R. (2000). "Saliency, stigma and standard," in *The Development of Standard English 1300-1800: Theories, Descriptions, Conflicts*, ed L. Wright, (London: Cambridge University Press). 57–72.
- Itti, L., and Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Res.* 49, 1295–1306. doi: 10.1016/j.visres.2008.09.007
- Jaeger, T. F., and Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition* 127, 57–83. doi: 10.1016/j.cognition.2012.10.013
- Kaschak, M. P., and Glenberg, A. M. (2004). This construction needs learned. *J. Exp. Psychol. Gen.* 133, 450–467. doi: 10.1037/0096-3445.133.3.450
- Kerswill, P., and Williams, A. (2002). "Saliency as an explanatory factor in language change: evidence from dialect levelling in urban England," in *Language Change: The Interplay of Internal, External and Extra-linguistic Factors*, eds M. C. Jones and E. Esch (Berlin: Mouton de Gruyter), 81–101. doi: 10.1515/9783110892598.81
- Kleinschmidt, D. F., and Jaeger, T. F. (2015). Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychol. Rev.* 122, 148–203. doi: 10.1037/a0038695
- Kraljic, T., Brennan, S. E., and Samuel, A. G. (2008). Accommodating variation: dialects, idiolects, and speech processing. *Cognition* 107, 54–81. doi: 10.1016/j.cognition.2007.07.013
- Kuhl, P. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Percept. Psychophys.* 50, 93–107. doi: 10.3758/BF03212211
- Kuperberg, G. R., and Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Lang. Cogn. Neurosci.* 31, 32–59. doi: 10.1080/23273798.2015.1102299
- Labov, W. (1966). *Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic Patterns*. Oxford: Blackwell.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106, 1126–1177. doi: 10.1016/j.cognition.2007.05.006
- Linzen, T., and Jaeger, T. F. (2016). Uncertainty and expectation in sentence processing: evidence from subcategorization distributions. *Cogn. Sci.* doi: 10.1111/cogs.12274. [Epub ahead of print].
- MacDonald, M. C., Pearlmutter, N., and Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703. doi: 10.1037/0033-295X.101.4.676
- McDonald, S. A., and Shillcock, C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychol. Sci.* 14, 648–652. doi: 10.1046/j.0956-7976.2003.psci_1480.x
- McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) on on-line sentence comprehension. *J. Mem. Lang.* 38, 283–312. doi: 10.1006/jmla.1997.2543
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *J. Lang. Soc. Psychol.* 18, 62–85. doi: 10.1177/0261927X99018001005
- Preston, D. R. (1996). Whaddayaknow?: The modes of folk linguistic awareness. *Lang. Awareness* 5, 40–74. doi: 10.1080/09658416.1996.9959890
- Qian, T., Jaeger, T. F., and Aslin, R. (2012). Learning to represent a multi-context environment: more than detecting changes. *Front. Psychol.* 3:228. doi: 10.3389/fpsyg.2012.00228
- Rácz, P. (2012). Operationalising saliency: Definite article reduction in the North of England. *Engl. Lang. Linguist.* 16, 57–79. doi: 10.1017/S1360674311000281
- Rácz, P. (2013). *Saliency in Sociolinguistics*. Berlin/New York: De Gruyter Mouton.
- Schirmunski, V. (1930). Sprachgeschichte und Siedlungsmundarten. *Germanisch-Romanische Monatsschrift* 18, 113–122, 171–188.
- Schmid, H. J. (2007). "Entrenchment, saliency, and basic levels," in *The Oxford Handbook of Cognitive Linguistics*, eds D. Geeraerts and H. Cuyckens (New York, NY: OUP), 117–138.
- Schmid, H.-J., and Günther, F. (2016). Towards a unified socio-cognitive framework for saliency in language. *Front. Psychol.* 7:1110. doi: 10.3389/fpsyg.2016.01110
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 379–423, 623–656.

- Shaw, J. A., Best, C. B., Mulak, K. E., Docherty, G. J., Evans, B. G., Foulkes, P., et al. (2015). "Effects of short-term exposure to unfamiliar regional accents: Australians' categorization of London and Yorkshire English consonants," in *Proceedings of the 15th Australasian International Conference on Speech Science and Technology* (Christchurch), 3–5.
- Smith, N. J., and Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition* 128, 302–319. doi: 10.1016/j.cognition.2013.02.013
- Squires, L. (2014a). Knowledge, processing, evaluation: Testing the sociolinguistic perception of English subject-verb agreement variation. *J. Engl. Linguist.* 42, 144–172. doi: 10.1177/0075424214526057
- Squires, L. (2014b). *Social Differences in the Processing of Grammatical Variation*. Penn Working Papers in Linguistics 20. Available online at: <http://repository.upenn.edu/pwpl/vol20/iss2/20/>
- Squires, L. (2016). "Processing grammatical differences: Perceiving versus noticing," in *Awareness and Control in Sociolinguistic Research*, ed A. Babel (Cambridge, UK: Cambridge University Press), 80–103.
- Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. *J. Lang. Soc. Psychol.* 18, 86–100. doi: 10.1177/0261927X99018001006
- Trudgill, P. (1986). *Dialects in Contact*, Oxford: Blackwell.
- Walker, A., and Hay, J. (2011). Congruence between 'word age' and 'voice age' facilitates lexical access. *Lab. Phonol.* 2, 219–237. doi: 10.1515/labphon.2011.007
- Weatherholtz, K., and Jaeger, T. F. (in press). Speech perception and generalization across talkers and accents. *Oxf. Res. Encycl. Linguist.*
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- The reviewer DM and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.
- Copyright © 2016 Jaeger and Weatherholtz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.