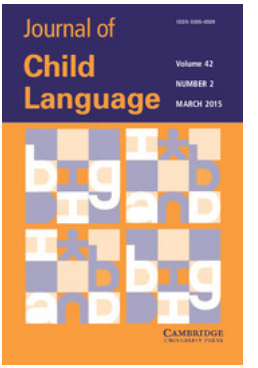

For and against frequencies

CHARLES YANG

Journal of Child Language / Volume 42 / Issue 02 / March 2015, pp 287 - 293
DOI: 10.1017/S0305000914000683, Published online: 03 February 2015

J. Child Lang. **42** (2015), 287–293. © Cambridge University Press 2015
doi:10.1017/S0305000914000683

For and against frequencies

CHARLES YANG
University of Pennsylvania

Frequency effects in child language are surely too sensible to dismiss. The first line in young Adam's transcript (Brown, 1973) is "play checkers", presumably because playing checkers is a more constant theme of life than, say, "program computers" or "burn leaves". These deserve more attention, as the authors (Ambridge et al., this issue) strike a balanced tone:

> We do not argue that sensitivity to input frequency must be the defining feature, or even the most important feature, of a successful account of acquisition (i.e., we do not argue for a frequency-DRIVEN or frequency-BASED mechanism). It is not difficult to think of factors that are more important than input frequency in at least some scenarios ... We argue, instead, for a learning mechanism that is minimally frequency SENSITIVE, under which input frequency need not be the chief determinant of acquisition in all cases.

It is hard to disagree with this modest proposal; I have made similar suggestions myself (e.g., Yang, 1999, 2002). And the authors' acknowledgment elsewhere in the target paper that frequency effects can be captured by a generative/nativist theory represents a revision of the usage-based tradition (Tomasello, 2003), one which casts generative grammar as inherently incompatible with the gradual and frequency-sensitive aspect of development.

Yet not all is well in the current compromise. In this brief commentary, I first examine the empirical basis of the frequency effects put forward by the authors. Turning then to some of the most intensively studied topics in child language, I wish to remind the reader why frequency effects, FOR VERY GOOD REASONS, have not always been at the forefront of acquisition research. I will conclude with some general remarks on the authors' theoretical framework to account for the effects of input frequency.

FREQUENCY EFFECTS?

The "ubiquity" of frequency effects, the authors claim, can be observed from perception to cognition, from music to language. That is all true. While I do not need an experiment to know that I recognize my daughter's face faster than that of a new student, it does not follow that such (rather obvious) frequency effects are of any theoretical importance. In the study of face perception, for example, the frequency effects cited by the authors

(e.g., Grammer & Thornhill, 1994) hardly get a mention in the current cognitive neuroscience literature. Rather more energy is devoted to theoretical questions such as whether face perception represents a domain-specific process in the mind/brain (Kanwisher, 2000), quite similar to the debate in language acquisition. Closer to home, the authors cite the role of frequency in language change (Bybee, 2010). Again, no one disputes the role of frequency, which has long been recognized as a factor in lexicalized changes (Labov, 1981; Kiparsky, 1995), as well as phonological processes such as lenition (Phillips, 1984). But by far most dominant force in language change is the structural conditioning of linguistic processes, whose effects sweep through the lexicon oblivious of frequency variation. This traditional (neo-grammarian) perspective now receives ample support in the modern quantitative studies of language change (Labov, 1994). For instance, in a study of the fronting of the diphthong nuclei /uw/, /ow/, and /aw/ in American English, Labov (2006) uses the dialect survey data (Labov, Ash & Boberg, 2006) to show that high-frequency words are neither more nor less advanced in the sound change than low-frequency ones, and that virtually all variation can be accounted for by purely phonetic factors. Similar results are observed in a large-scale longitudinal corpus on current changes in progress (Labov, Rosenfelder & Fruehwald, 2013; Labov, 2014).

Indeed, while the frequency effects noted by the authors are generally well established, some claims deserve a closer inspection. I will limit myself to inflectional morphology and multiple word expressions, for which a re-examination can be carried out quickly with the data in the public domain. The authors use over-regularization errors to highlight the role of absolute frequency in acquisition: "the high-frequency irregulars *blew* and *feet* are less likely to be over-regularized (e.g., **blowed*, **foots*) than the low-frequency irregulars *drank* and *shelves* (e.g., **drinked* and **shelfs*)". But child language data shows otherwise. In a sample of approximately 3 million words of child English (MacWhinney, 2000), *blew* was regularized 17 out of 41 times (41%), while *drank* was regularized 13 out of 39 times (33%), and *feet* was regularized 13 out of 393 times (3%), but every use of *shelf* in the plural was correct (albeit only 3 tokens). When establishing the role of type frequency in generalization, the authors claim that "*say*→*said*" does not lead the child to say "*play*→**pled*"' or "*obey*→**obed*" because of its singleton type frequency, which is to be contrasted with patterns of higher type frequency (e.g., *blow/blew, know/knew, grow/grew, throw/threw*) that are more productive. Unfortunately, not a single instance of overextension of the *ow–ew* pattern (or *ay–ed*) can be found in the child data. Generalizations of irregular forms are virtually absent in child English (Xu & Pinker, 1995) as well as many other languages (Clahsen, Aveldo & Roca, 2002; Guasti, 2004): there is no evidence for varying

degrees of productivity among the irregular forms as a function of their (type) frequencies.

Then there are putative frequency effects that are not. Certain statistical patterns may appear to support one theoretical perspective but turn out to support quite the opposite once properly assessed. In the discussion of multiword and simple syntactic constructions, the authors draw attention to high-frequency combinations, which have been held as evidence for stored holophrases rather than the output of a genuinely productive system (Tomasello, 2000; cf. Lieven, Pine & Barnes, 1992). For instance, the expression 'give me'' (sometimes glossed as 'gimme') has been treated as an unanalyzed unit rather than formed as by a combinatorial rule. There is little doubt that 'give me' has significantly higher frequency that other 'give'–pronoun pairs. In the transcripts of Adam, Eve, and Sarah (Brown, Cazden & Bellugi, 1973), the frequencies of 'give me', 'give him', and 'give her'' are 95:15:12 (or 7·91:1·23:1). When a child's language is sampled, it would not be surprising if only 'give me' can be found in the transcript. But we still need a statistical baseline to determine if the sparsely represented multiword combinations are the effect of holistic storage or syntactic combination. As such, note that the frequency ratios for 'me', 'him', and 'her' from the same dataset are 2870:466:364 (or 7·88:1·28:1), which are almost completely in parallel with the verb–pronoun pairs. The frequency effects of these expressions, then, can be attributed entirely to the marginal probabilities of the pronouns, providing evidence for a categorical rule rather than holistic storage. In other lines of work, I have considered the nature of frequency effects in the syntax of noun phrases and verb phrases (Kowalski & Yang, 2012; Yang, 2013; see also Valian, Solt & Stewart, 2009). Frequency effects may be apparent, but their interpretation requires more precisely formulated predictions of competing theoretical approaches.

ANTI-FREQUENCY EFFECTS

One of the goals of the target paper is to convince 'the most hardened classicist' that frequency matters. They have succeeded, but it is also worth pondering why the classicists, presumably the generativists, are so hardened. The reason, I believe, is that researchers have traditionally, and with ample justification, focused on empirical cases where frequency does not, and probably cannot, play any significant role.

It is generally assumed that a theory of language and language acquisition should provide a broad account for the speaker's linguistic knowledge. On the one hand, this includes expressions language users are capable of producing (and sometimes do produce), and on the other, expressions that have not, and in fact cannot, be produced, for they are prohibited by

universal or language-specific constraints. Examples such as structural dependence, island conditions, and Binding Theory immediately spring to mind, and there have been robust confirmations about their grammaticality judgment in both children and adults; see Crain and Thorton (2000), Sprouse (2011), and Sprouse and Almeida (2012) for summaries. These aspects of language acquisition have been a major focus in the generative tradition but has not been frequently discussed in the usage-based literature or the present paper. To account for the things speakers cannot say, the role of input frequency seems either minimal or insufficient: ungrammatical forms would rarely if ever appear in the input, and crucial disconfirming data may not be robustly represented to be useful to the learner (Legate & Yang, 2002).

Even in cases where input clearly matters, its role is not always transparent in children's language use. "Play checker" does show frequency effects, but Adam taught us a good deal more about language acquisition with "Where go?", "he go out", "He should have holded his horses", and other innovations that stubbornly defy the input. It is difficult to think of more prominent or productive research topics than the null-subject phenomenon (Bloom, 1970), the Optional Infinitive stage (Wexler, 1994), and the past tense debate (Rumelhart & McClelland, 1986; Pinker & Prince, 1988). All three cases, which have generated a large body of cross-linguistic literature, concern robust usage patterns that quantitatively, and often qualitatively, deviate from the input data; perhaps that is why researchers have found them to be so interesting. In all three cases, the child gradually eliminates the errors and converges on to the target form, which suggests – here we are in agreement with the authors – a role for frequency-sensitive learning processes (Yang, 2002; Legate & Yang, 2007). But the effect of the input must be indirect.[1] After all, adults do not generally omit subjects and objects, so any instance of null subjects and objects constitutes evidence AGAINST frequency effects. To complicate the matter further, there have been many cross-linguistic studies on subject/object omission, which appear paradoxical under the typical interpretation of frequency effects. While the acquisition of subject/object use is famously delayed in English and other obligatory subject languages, children acquiring pro-drop (e.g., Italian) and topic-drop (e.g., Chinese, Korean) languages are much closer to adult level usage at a much earlier age (Valian, 1991; Wang, Lillo-Martin, Best & Levitt, 1992; Kim, 2000).

[1] Earlier we noted the input frequency and child performance discrepancies in the acquisition of irregular past tense. The breakdown of frequency effects, which poses difficulty for the associative memory account of past tense learning, turns out to support the traditional view (Bloch, 1947) that irregular verbs are organized into classes that are defined by lexicalized (i.e., unproductive) rules (Yang, 2002).

It would be awkward that the English-learning child be less capable of matching the probabilities of language use: sociolinguists have known for a long time that children are exceptionally good at learning the statistical distribution of linguistic forms (Labov, 1989; Roberts & Labov, 1995; Smith, Durham & Fortune, 2009). A straightforward account (see Yang, 2004, for a summary) is to assume that acquisition follows a selectionist scheme in which the learner chooses the options employed in her language ([±pro drop], [±topic drop]) on the basis of the input, and that differences in the acquisition of various languages are due to the quantitative differences in the amount of disambiguating evidence in the input. As the authors note, Universal Grammar is consistent with frequency effects.[2]

INTERACTIONS OR OBFUSCATIONS?

My last critical comment is more conceptual and theoretical. It is not clear to me what mode of explanation the authors are committed to, since it is very difficult to pin down any specific predictions of frequency effects. For example, it is not clear what "a frequency-DRIVEN or frequency-BASED mechanism" actually is, which the authors are reluctant to commit to but which features in most of the examples where the input frequency is transparently reflected in child language use.

The overall framework, dubbed the interaction thesis, is that all aspects of language (as well as non-linguistic aspects of cognition such as memory) are jointly engaged in language acquisition. This surely covers all the bases, but how are these factors supposed to interact? Are some interactions more important than others, and if so, when? In their account of frequency effects, the authors invoke factors that operate as conflicting forces from the entire interactive spectrum. There are effects concerning abstract categories as well as concrete lexical entries, frequencies can be absolute but also relative, sometimes pertaining to types while sometimes to tokens, errors can be caused by frequency as well as prevented, etc. When frequency effects fail to turn up, that just proves the INVERSE FREQUENCY EFFECTS. Heads I win, tails you lose. There is little doubt that frequency matters, and we are grateful to the authors for the reminder. But burying frequency in an entangled web of interactions does little to elucidate how input data is integrated into the development of language.

[2] It is not fair for the authors to regard the parameter setting model of Sakas and Fodor (2012) as inherently inconsistent with frequency effects. The key idea in that work is that the learner tends to structural units ("treelets"), rather than word strings, in disambiguating parameter values. So far as I can tell, the model can trivially incorporate frequency information: the setting of the parameter values will be sensitive to the frequency of treelets the learner encounters in the input.

REFERENCES

Bloch, B. (1947). English verb inflection. *Language* **23**(4), 399–418.

Bloom, L. (1970). *Language development: form and function in emerging grammar*. Cambridge, MA: MIT Press.

Brown, R. (1973). A first language: the early stages. Cambridge, MA: Harvard University Press.

Brown, R., Cazden, C. B. & Bellugi, U. (1973). The child's grammar from I to III. In C. A. Ferguson & D. I. Slobin (eds), *Studies of child language development* (pp. 295–333). New York: Holt, Rinehart and Winston.

Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Clahsen, H., Aveledo, F., and Roca, I. (2002). The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language* **29**(3), 591–622.

Crain, S. & Thornton, R. (2000). *Investigations in universal grammar: a guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Grammer, K. & Thornhill, R. (1994). Human (homo sapiens) facial attractiveness and sexual selection: the role of symmetry and averageness. *Journal of Comparative Psychology* **108**(3), 233–242.

Guasti, M. T. (2004). *Language acquisition: the growth of grammar*. Cambridge, MA: MIT Press.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience* **3**, 759–763.

Kim, Y.-J. (2000). Subject/object drop in the acquisition of Korean: a cross-linguistic comparison. *Journal of East Asian Linguistics* **9**(4), 325–351.

Kiparsky, P. (1995). The phonological basis of sound change. In J. Goldsmith (ed.), *The handbook of phonological theory* (pp.640–670). Cambridge: Blackwell.

Kowalski, A. & Yang, C. (2012). Verb islands in child and adult language. In A. K. Biller, E. Y. Chung & A. E. Kimball (eds), *BUCLD 36: Proceedings of the 36th annual Boston University Conference on Language Development*, vol. 1 (pp. 281–289). Somerville, MA: Cascadilla Press.

Labov, W. (1981). Resolving the neogrammarian controversy. *Language* **57**(?), 267–308.

Labov, W. (1989). The child as linguistic historian. *Language Variation and Change* **1**(1), 85–97.

Labov, W. (1994). *Principles of linguistic change: internal factors*. Oxford: Blackwell.

Labov, W. (2006). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics* **34**(4), 500–515.

Labov, W. (2014). The regularity of regular sound change. Unpublished manuscript, University of Pennsylvania.

Labov, W., Ash, S. & Boberg, C. (2006). *The atlas of North American English: phonetics, phonology, and sound change*. Berlin: Mouton de Gruyter.

Labov, W., Rosenfelder, I. & Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: linear incrementation, reversal and re-analysis. *Language* **89**(1), 30–65.

Legate, J. A. & Yang, C. (2002). Empirical re-assessment of stimulus poverty arguments. *Linguistic Review* **19**, 151–162.

Legate, J. A. & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition* **14**(3), 315–344.

Lieven, E. V., Pine, J. M. & Barnes, H. D. (1992). Individual differences in early vocabulary development: redefining the referential–expressive distinction. *Journal of Child Language* **19** (2), 287–310.

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk*, 3rd ed. Mahwah, NJ: Lawrence Erlbaum.

Phillips, B. (1984). Word frequency and the actuation of sound change. *Language* **60**(2), 320–342.

Pinker, S. & Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* **28**(1), 73–193.

Roberts, J. & Labov, W. (1995). Learning to talk Philadelphian: acquisition of short *a* by preschool children. *Language Variation and Change* **7**, 101–112.

Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart & the PDP Research Group (eds), *Parallel distributed processing: explorations into the microstructure of cognition. Volume 2: Psychological and biological models*, 216–271. Cambridge, MA: MIT Press.
Sakas, W. G. & Fodor, J. D. (2012). Disambiguating syntactic triggers. *Language Acquisition* **19**(2), 83–143.
Smith, J., Durham, M. & Fortune, L. (2009). Universal and dialect-specific pathways of acquisition: caregivers, children, and t/d deletion. *Language Variation and Change* **21**(1), 69–95.
Sprouse, J. (2011). A validation of amazon mechanical turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods* **43**(1), 155–167.
Sprouse, J. & Almeida, D. (2012). Assessing the reliability of textbook data in syntax: Adger's *Core Syntax*. *Journal of Linguistics* **48**(3), 609–652.
Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* **11**(1/2), 61–82.
Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition* **40**(1/2), 21–81.
Valian, V., Solt, S. & Stewart, J. (2009). Abstract categories or limited-scope formulae? The case of children's determiners. *Journal of Child Language* **36**(4), 743–778.
Wang, Q., Lillo-Martin, D., Best, C. T. & Levitt, A. (1992). Null subject versus null object: some evidence from the acquisition of Chinese and English. *Language Acquisition* **2**(3), 221–254.
Wexler, K. (1994). Optional infinitives, verb movement and the economy of derivation in child grammar. In D. Lightfoot & N. Hornstein (eds), *Verb movement*, 305–350. Cambridge: Cambridge University Press.
Xu, F. & Pinker, S. (1995). Weird past tense forms. *Journal of Child Language* **22**(3), 531–556.
Yang, C. (1999). A selectionist theory of language development. In Robert Dale & Kenneth W. Church (Eds.), *Proceedings of ACL* (pp. 429–435). Stroudsburg, PA: Associaton for Computational Linguistics.
Yang, C. (2002). *Knowledge and learning in natural language*. Oxford: Oxford University Press.
Yang, C. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences* **8**(10), 451–456.
Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences* **110**(16), 6324–6327.