Uninterpretable features: psychology and plasticity in second language learnability

David Birdsong

Second Language Research 2009 25: 235
DOI: 10.1177/0267658308100285

Second Language Research 25,2 (2009); pp. 235–243

Uninterpretable features: psychology and plasticity in second language learnability

David Birdsong *University of Texas*

Received June 2008; revised September 2008; accepted October 2008

This commentary addresses the relevance of detectability to a theory of learning uninterpretable features in the second language (L2). Detectability of features is illustrated in an application of Signal Detection Theory. By analogy with development of phonemic categories in the first language (L1), the notion of paring down the repertoire of uninterpretable features is considered.

Keywords: uninterpretable feature, learnability, signal detection, second language acquisition

I Introduction

Lardiere asserts that, at the level of the feature, the first language (L1) mediates second language (L2) learning structurally and procedurally: 'learners use L1 feature configurations as a departure point for what to look for in the L2.' Within this context she connects two provocative claims: Learnability is a matter of detection, and there is no difference in inherent detectability (and therefore, no difference in inherent learnability) between interpretable and uninterpretable features. Quoting Lardiere (this issue: 214): 'any feature contrast that is detectable is, in principle, ultimately acquirable.' Further quoting: 'I do not see that the detectability of formal contrasts is likely to be split along interpretable vs. uninterpretable feature lines' (note 35).

Lardiere's position on ultimate learnability opposes that of Hawkins and colleagues (e.g. Hawkins and Hattori, 2006) and Tsimpli and colleagues (e.g. Tsimpli and Dimitrakopoulou, 2007). These researchers

Address for correspondence: David Birdsong, Department of French and Italian, University of Texas at Austin, 1 University Station B7600, Austin, TX 78712–0224, USA; email: birdsong@ccwf.cc.utexas.edu

 10.1177/0267658308100285

claim that the ability to correctly represent L2 uninterpretable features not selected by the L1 is lost after the closure of a critical period, while the ability to acquire interpretable features is not compromised. Tsimpli and Dimitrakopoulou (2007: 217) posit a role for detection in learnability, though not as emphatically as Lardiere: 'the claim is that interpretable features are accessible to the L2 learner whereas uninterpretable features are difficult to identify and analyse in the L2 input due to persistent, maturationally-based, L1 effects on adult L2 grammars.'

Fundamental questions loom. What is meant by detecting? How is information from various sources (pragmatics, information structure, prosody, L1, UG, etc.) integrated in detection? Is detection required for learning? Lardiere justifiably leaves these daunting topics for others to pursue. Consequently my initial inclination is to not hold Lardiere's feet to the fire, preferring instead to read 'detecting' as a loose shorthand for as-yet-unspecified mental processes, domain-general in nature (though guided by knowledge given by the L1 or UG), which are applied in the local domain of feature acquisition (and which await integration into formal models of L2 acquisition and knowledge). Indeed, a strict reading might force one into re-re-consideration of Krashen's Monitor Model, Schmidt's conscious and unconscious noticing, Slobin's Operating Principles, Van Patten's focus on form, and so on (Slobin, 1973; Krashen, 1982; Schmidt, 1993; VanPatten, 2004). One would have to rule in or out the applicability to L2 feature acquisition of over-regularization, processability, and Bayesian-probabilistic and statistical learning (Aslin *et al.*, 1998; Ellis, 2002). And to do justice to the issue of locus of searching-and-detecting one would have to reconcile top-down with bottom-up theories of identifying the problem space (i.e. 'guided searching'), and accommodate noticing of both superficial and abstract linguistic structure. So the temptation is to not 'go there'.[1]

But it is hard to ignore Lardiere's repeated references to looking for, noticing, and detecting. Moreover, I agree with Lardiere that the psychology of detection in L2 learning is inherently interesting (note 35). It is all the more interesting given Lardiere's expressed premise that learners' noticing of L1–L2 featural differences *à la* contrastive analysis is well adapted to L2 acquisition. Finally, we have to wonder in

[1] I refer readers to authorities who do 'go there'. Carroll (2001) and Truscott and Sharwood Smith (2004) consider at length the linkage of processing mechanisms to formal theories of L2 learning. I am grateful to Susanne Carroll for discussion of issues raised in this commentary.

what functional sense are uninterpretable features thought to be no less detectable – and thereby no less learnable – than interpretable features.

II Detecting feature contrasts

As a way into these issues, consider the principles of Signal Detection Theory (SDT; Green and Swets, 1966) as they might apply to the learning of L2 features. In SDT, the actual detection of a signal (a telephone ringing as the stereo blares, a particular phoneme within a sequence of sounds, an article or gender marker inside a phrase) depends on three factors: the ratio of signal to noise, the expectations of the processor, and the motivation to make a decision about the presence or absence of a signal.

- Signal-to-noise ratio: This mathematical relationship obtaining between the strength of the signal and the strength of the background noise characterizes the inherent detectability of the signal. In the context of acoustics this could be as simple a calculation as dividing the decibels of the signal by the decibels of the noise.
- Expectation: A person who knows a phone call is imminent will be attentive cognitively and attuned auditorily to the sound of a phone ringing. Thus, situational, expectancy-based attention and acuity modulate the detection of a signal.
- Motivation: This factor speaks to the pay-off for detection. For example, even when asleep in the middle of the night, parents are supremely motivated to detect a choking sound from their baby, and the obvious pay-off is the infant's welfare.

Taking this example to suggest how the three factors might interact, we can imagine that the signal-dampening effects of 'noise' (here, the mental state of sleep) would be attenuated by heightened attentiveness to sounds coming from the baby. Heightened attentiveness is seen in turn as deriving from heightened motivation to detect.

With all due disclaimers regarding oversimplification and speculativeness, SDT (or some functional variant) could be applied to detection of feature contrasts:

- Signal-to-noise ratio: The contrast between the singular noun *row* and the plural noun *rows* is given by the plural affix /z/. It is safe to assume that the acoustics of the /z/ signal does not differ between

/z/-final singulars (*rose*) and /z/-final plurals (*rows*). That is, the signal strength of plural /z/ should be no less robust than any other word-final /z/. Holding noise constant across /z/-finals, the inherent detectability of the feature contrast is not in question.

- Expectation: As Lardiere states, in adult second language acquisition, expectations for linguistic form are derived from the L1: in the default case, feature contrasts in the L2 are initially expected to be isomorphic and co-extensive with feature contrasts of the L1. Other sources of expectation would be language universals, intuitions for markedness and psychotypological distance, and any presumed UG-given constraints on linguistic form that are accessible in adult language learning.
- Motivation: Following Lardiere's logic, the pay-off for detection of a given feature contrast is acquisition of the feature. An individual L2 learner might dedicate only the attention required for understanding or being understood at a chosen level of functioning. Or a given learner might be motivated to pass for a native, and in order to do so would attend to all inherently detectable elements of the target language.

Returning to Lardiere's position that what is detectable is learnable, is there any reason to believe that uninterpretable features are not detectable? Or, perhaps more accurately, are uninterpretable feature contrasts less detectable than interpretable feature contrasts?[2] Like many of the issues raised here, this is ultimately an empirical question. One approach would be to demonstrate that the 'signal' of uninterpretable features is relatively weak. Consider the possibility that, for various reasons, uninterpretable features might be more subject to phonological reduction or elision than are interpretable features. For example, the uninterpretable person and number feature contrasts for English present tense *be* (*am*, *are*, *is*) are often elided to *'m*, *'re*, *'s*.

[2] A distinction is to be made here between inherent detectability (operationalized in terms of the signal-to-noise ratio) and actual detection. Assuming a linear relationship, the strength of former predicts the frequency (i.e. states the likelihood) of the latter. The relationship may not be linear, however: there may be threshold effects, for example. This said, one might safely claim that – as a function of signal-to-noise ratios – some signals are more likely to be detected than others. I therefore continue to refer to features and their associated contrasts in terms of relative detectability. Note that detection – an output of signal processing – is understood in an absolute sense: a feature contrast is either detected or not. That is, detection is an effect that is predicted by the variable of signal-to-noise ratio and modulated by expectations and motivation.

But elision and reduction are not exclusive to uninterpretable features. For example, in normally-paced spoken English the Boolean operators 'and' and 'or' are routinely reduced and 'not' is routinely elided. (Certainly, the ultimate learnability of these terms is not compromised by signal weakness.) Still, it is conceivable that, in the aggregate and for a given language, formal contrasts in uninterpretable features are subject to more frequent or more significant signal reductions than are interpretable feature contrasts. It is likewise possible that, with respect to potential prominence in information structure (e.g. susceptibility to focus accent), uninterpretable feature contrasts are relatively weak in the aggregate.

Rather than belabour these issues, I simply offer a few observations:

- Signal-strength differences between interpretable and uninterpretable feature contrasts are conceivable; within each feature class, differences in signal strength would be likely.
- If L2 feature learning is keyed to detectability, then hypothetically there should be correlational evidence of this association. Further, we would want to know if significant correlations obtain across both interpretable and uninterpretable feature contrasts, or only within one or the other class of feature.
- As Lardiere points out (note 35) detection of feature contrasts may be aided by literacy. In self-paced reading, attention and analysis can be allocated at will to anything on the printed page. Functionally, literacy should level the playing field, neutralizing any putative role of differential detectability in the learning of interpretable and uninterpretable features. The affordances of literacy would presumably privilege literates over non-literates in the acquisition of both interpretable and uninterpretable features.

III Plasticity and uninterpretable features

It is well known that infants past the age of six months become progressively unable to discriminate between language sounds that are not in their established phonetic repertoire. It is less well known that performance declines are to a large extent reversible. This finding is replicated in refined behavioural techniques and on brain-based measures such as event-related potentials. Moreover, adults can be trained to hear phonic distinctions that they are supposed to be 'deaf' to, as in the case of discrimination of English /r/–/l/ by native speakers of Japanese. Nativelike

performance on various measures is seen in highly proficient L2 users, particularly those for whom the L2 is dominant. Werker and Tees (2005) review these and related findings. They conclude that perceptual flexibility is not permanently attenuated with advancing age. They suggest that developmental changes in discrimination are not the result of a loss of sensitivity, but reflect ongoing representational organization in response to language input. However functionally consolidated this organization may be at a given time, the perceptual mechanism maintains potential responsiveness to linguistic input.

By analogy to sound discrimination, we may ask: Do L2 learners become progressively unable to recognize differences between the formal features of their L1 and the formal features of their L2? Are observed deficits reversible? Are deficits revealed in some experimental methods but not others? Is there a role for training? Can highly proficient L2 learners and L2 dominants perform like natives in the domain of feature discrimination (and feature learning)?

Pursuing this analogy, Lardiere elaborates on the first two questions. Following Brown (2000), Lardiere (this issue: 213) cites the example of voice contrasts that distinguish phonemes. If voicing is represented 'somewhere' in the L1 phonology of the adult L2 learner, it can be recruited to enable discrimination of any non-native contrast that involves voicing. Once this feature is engaged in the discrimination process, it is available as a compositional element of a new phonic category. By analogy, if a given interpretable morphosyntactic feature is available anywhere in the grammar of the L1, then it potentially enables discrimination of relevant semantic/syntactic contrasts in the grammar. This morphosyntactic feature is then available as a 'building block' for composing and reconfiguring lexical items.

Lardiere's quotation from Brown (2000: 20) makes it clear that the phoneme categories (i.e. analogues of composed lexical items in morphosyntax) of the L1 come into play in L2 acquisition: incoming sounds are initially sorted into these categories. As initial-state representations, they constrain and guide perception. However, their effects may not endure past early stages of acquisition; the subsequent establishment of new phonological categories is not ruled out.

Are unfamiliar morphosyntactic contrasts in the L2 hard to 'hear' at first? Yes. Are the L1-/maturationally-constrained biases underlying this difficulty permanent (and perhaps across the board)? No. For proponents of representational deficit views, however, the answer to

both questions is Yes, the biases make it difficult to establish L2-like representations initially, and they are permanent impediments to learning if the L2 learner is dealing with uninterpretable features unselected in the L1. On this view, by virtue of experience in acquiring and continuously using the L1, the class of uninterpretable features is pared down to contain only those represented in the L1.

Werker and Tees (2005) note that early evidence suggesting persistence of the auditory-discrimination deficit was limited, and that as more experimentation with procedural controls and sensitive methodologies came on line, the newer evidence revealed that sensitivity to acoustic signals is not completely and irreversibly narrowed. This is not the place to critically review evidence from the experimental literature that relates in one way or another to adult L2 acquisition of uninterpretable features. However, it is fair to say that the findings are mixed. Results vary by task (e.g. on-line vs. off-line tasks) and by measure (e.g. behavioral vs. brain-based measures). Investigation of subject biographical factors has not been systematic; however, it is likely that results vary by subjects' age of arrival in the L2 context, L2 proficiency, relative L1 vs. L2 use, and relative L1 vs. L2 dominance. Connecting such subject variables with outcome measures would be revealing. Also potentially instructive would be comparisons of outcomes (on the same tasks and measures, with the same subjects) for different uninterpretable features. Here the relevant question is: Are all uninterpretable features similarly (un)learnable?

Werker and Tees take seriously the question of ultimate potential alongside transient and partial deficits. To this end, the researchers identify a role for training, for example using exaggerated acoustic cues in order to enable discrimination. In the case of uninterpretable features, we cannot exclude the possibility that training can be optimized (in terms of implicit/explicit input conditions, presence/absence of positive/negative feedback, drawing attention to L1–L2 differences, instruction in 'detection' of unselected uninterpretable features, etc.) for effective L2 learning. Similarly, the role of benign circumstances in naturalistic settings is not to be overlooked. If the idea is to compare observed end-state attainment in the L1 with observed end-state attainment in the L2, then we should not deny to L2 learners the external conditions that are known to favour language acquisition, starting with massive linguistic input and interaction with natives.

Subjects under other facilitating conditions might be studied, including L2-dominant adults of two varieties: those whose routine L2 use

exceeds L1 use, and those whose psycholinguistic processing abilities in the L2 exceed those in the L1. In the extreme, the use of and contact with the mature L1 would be eliminated altogether, as in the case of the Korean adoptees studied by Pallier *et al.* (2003).

We return then to the notion that, with maturation, uninterpretable features not selected by the L1 become unavailable, and cannot be acquired by adult L2 learners. In its strong form, the elements of this position are that all unselected uninterpretable features are unlearnable; that essentially all adult learners, including L2-dominants, fail to learn unselected uninterpretable features; that failure to learn can be revealed across multiple sensitive tasks and behavioral/imaging measures; and that there is nothing one can do (in terms of input and training) to enable the learning of unselected uninterpretable features. It is consistent with findings reported by Werker and Tees (2005) – and with the spirit of Lardiere's remarks on detecting uninterpretable features – to question each of the components of this claim.

IV References

Aslin, R.N., Saffran, J.R. and **Newport, E.L.** 1998: Computation of conditional probability statistics by human infants. *Psychological Science* 9, 321–24.

Brown, C. 2000: The interrelation between speech perception and phonological acquisition from infant to adult. In Archibald J., editor, *Second language acquisition and linguistic theory*. Malden, MA: Blackwell, 4–63.

Carroll, S.E. 2001: *Input and evidence: the raw material of second language acquisition.* Amsterdam: John Benjamins.

Ellis, N.C. 2002: Frequency effects in language acquisition: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24, 143–88.

Green, D.M. and **Swets, J.A.** 1966: *Signal detection theory and psychophysics.* New York: Wiley.

Hawkins, R. and **Hattori, H.** 2006: Interpretation of English multiple *wh*-questions by Japanese speakers: a missing uninterpretable feature account. *Second Language Research* 22, 269–301.

Krashen, S.D. 1982: *Principles and practice in second language acquisition*. Oxford: Pergamon.

Pallier, C., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E. and **Mehler, J.** 2003: Brain imaging of language plasticity in adopted adults: can a second language replace the first? *Cerebral Cortex* 13, 155–161.

Schmidt, R. 1993: Awareness and second language acquisition. *Annual Review of Applied Linguistics* 13, 206–26.

Slobin, D.I. 1973: Cognitive prerequisites for the development of grammar. In Ferguson, C.A. and Slobin, D.I., editors, *Studies of child language development*. New York: Holt, Rinehart and Winston, 175–208.

Truscott, J. and **Sharwood Smith, M.** 2004: Acquisition by processing: a modular perspective on language development. *Bilingualism: Language and Cognition* 7, 1–20.

Tsimpli, I.-M. and **Dimitrakopoulou, M.** 2007: The Interpretability Hypothesis: evidence from *wh*-interrogatives in second language acquisition. *Second Language Research* 23, 215–42.

VanPatten, B. 2004: Input processing in second language acquisition. In VanPatten, B., editor, *Processing instruction: theory, research, and commentary*. Mahwah, NJ: Lawrence Erlbaum, 5–31.

Werker, J.F. and **Tees, R.C.** 2005: Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology* 46, 233–51.