

Peter de Swart

A (single) case for heritage speakers?

Keywords: heritage languages, individual differences, literacy, educational linguistics, incomplete acquisition

Peter de Swart: Department of Linguistics, Radboud University Nijmegen.

E-mail: p.deswart@let.ru.nl

1 The more the merrier

Present-day linguists can choose from a wide variety of methods to establish facts about language and language users, ranging from introspections to corpus studies and experiments, to name but a few. At the same time, the linguist has to decide from which population of speakers he wants to extract his data. These include adult ('monolingual') native speakers, children acquiring a native language, speakers learning a second language or speakers showing some language-related pathology. Benmamoun, Montrul and Polinsky's (BMP) aim is to bring so-called heritage speakers to the attention of (theoretical) linguists as a new population to be added to this list. In their opinion heritage speakers are a valuable source of data that can 'feed back into linguistic theory and help to promote its progress' (p. 154).

When it comes to understanding language, 'the more the merrier' should be our motto. Only if we take into account multiple sources of information and look at phenomena from different perspectives, we can hope to unravel the language puzzle (Wasow and Arnold 2005). BMP's plea to consider heritage speakers as a source of information can therefore only be endorsed. The question remains what can be learned from the data of heritage speakers. Can we indeed use this data to promote progress of linguistic theory? Does this data have a bearing on theoretical discussions? The answer, at this stage, is not immediately clear. The examples put forward by the authors are not extremely convincing and if these represent the strongest cases, which is hard to judge as an outsider to the field of heritage linguistics, than my answer has to be negative (for now). Heritage linguistics, as represented in this target article, seems not yet in a position to firmly support theoretical generalizations. This is mainly due to methodological issues, in particular the high intersubject variability present in studies with heritage speakers. Below, I will first discuss two of the examples with assumed theoretical

relevance, that of the universality of the noun-verb distinction and the distinction between structural and inherent case, before going into the methodological issues.

1.1 Nouns vs. verbs

BMP provide the universality of the noun-verb distinction as an example where data from heritage speakers may provide a valuable contribution to theory construction. The noun-verb distinction has been argued to be a fundamental, universal property of language, even though some languages are said to lack it. Studies with heritage speakers of Russian and Korean have shown that these speakers recognize verbal items more accurately and faster than nominal (and adjectival) items. This suggests indeed that the noun-verb distinction is retained by these heritage speakers. But these results do by no means ‘offer new support for the universality of the noun-verb distinction’ (p. 148). Given that heritage speakers have had prior exposure to the distinction between nouns and verbs, either in their first or their second language (or in both), it is no surprise that they distinguish between them. It would be more surprising if they had not distinguished between the two categories. In this respect, data from heritage speakers does not seem to add anything that we cannot learn from other populations. We would also not conclude from a similar finding in the first language of monolingual or bilingual speakers that the noun-verb distinction is a universal. Evidence for the universality of this distinction can only be found through the study of grammars of individual languages and their comparison (e.g. van Lier 2009), children learning their first language (e.g. Braine 1987) or from neurolinguistic studies showing that nouns and verbs are located in different regions of the brain (e.g. Corina et al. 2005)

The relevant question is why heritage speakers are faster and better at recognizing verbs. The suggestion by the authors that this may be due to the smaller class size for verbs sounds plausible, as this would considerably narrow down the search space for verbs. This proposal does not only affect heritage speakers, and we would expect to find similar effects in other populations.

1.2 Structural vs. inherent case

Linguists of different theoretical persuasions agree that some cases have a more structural use, coding grammatical functions such as subject and object, whereas

others are more closely linked to semantic information (Haspelmath 2009). In generative grammar this distinction is known as that between structural and inherent case. Which case belongs to which category is, however, often still a matter of debate. BMP note that in practice, researchers ‘tend to impose the structural/inherent distinction top-down, by analogy with more familiar languages, and then form expectations based on those familiar languages’ (p. 155). I agree with the authors that such reasoning from a familiar situation into the unknown is ‘not always a successful strategy’ (p. 155), even though it is more widely applied (see Haspelmath 2010).

Can data from heritage speakers provide us with insight into the theoretical discussion concerning structural and inherent case? According to the authors it can, and their discussion about the status of ergative case (*-ne*) in Hindi provides an example. Based on studies of the use of case marking in heritage speakers of Russian (and Spanish), BMP come to the following generalization (their (25)): ‘In a heritage language, STRUCTURAL CASE of the baseline is replaced by an unmarked case, whereas INHERENT CASE is maintained (although its morphological exponent may change compared to the baseline)’ (p. 160). A similar study on the use of case marking by heritage speakers of Hindi shows that these speakers exhibit a ‘significant erosion’ (p. 161) of ergative case marking. The authors maintain that this suggests that the Hindi ergative is a structural case.

This conclusion may turn out to be the right one in the end, but as of yet is not firmly grounded. First, it is established on the basis of the same kind of reasoning from familiar to unfamiliar situations that the authors deemed ‘not always a successful strategy’. A better strategy would be to make a language-internal comparison to find commonalities in patterns with cases that are unmistakably of one type or the other. Still, there may be other reasons that can account for the erosion of a certain case (see below). Second, the generalization from which the reasoning starts is not as firm as it may seem. It is established mainly on the results of the comprehension study presented in their Figure 2. These results show a clear split between structural accusative case and the other inherent cases. To conclude from this that the heritage speakers maintain inherent case is an overinterpretation of the significant difference which does not take into account the actual scores; something that is more often found in experimental studies. The scores for the inherent cases all fall in the middle of the scale (between 3 and 4) and (crucially) appear different from that of the native controls (although this is hard to tell from the graph provided, similar findings are reported in Montrul and Bowles 2009: 375). This suggests that heritage speakers are somewhat indeterminate with respect to the acceptability of omitting inherent cases. In addition, the patterns in the reported production data are much less

clear. Setting these aside opens the way to a crispier generalization but also invites discussion of the general relevance of production data (in studies on heritage grammars). The card of performance limitations should only be played when one has independent evidence for this, e.g. via comparison of speech rate in both languages and/or via comparison with behavior on similar phenomena in the dominant language.

In sum, the data presented by the authors does not convincingly add to the theoretical discussion, as envisioned by the authors. This is not to deny that data from heritage speakers can supply interesting data in this area.

2 Morphology as the ‘weakest link’ in heritage grammars

Even if the generalization that structural case is replaced and inherent case is maintained in the language of heritage speakers turns out to be correct, the more interesting question is why this is the case. The generalization is at odds with what we know from the behavior of these cases in other populations. Lamers and Ruigendijk (2009), for instance, report that inherent case is more vulnerable than structural case in spontaneous speech production of aphasic speakers. As long as the structural/inherent divide has not been settled, it seems more fruitful to narrow down the question to individual cases. Why are certain (uses of) cases preserved but others not? I will focus here on the use of differential case marking. Many studies have investigated the use of case marking by heritage speakers (e.g. Montrul and Bowles 2009, Montrul and Sánchez-Walker 2013, and Montrul et al. 2012). The pattern emerging from these data, simplifying to some extent (cf. BMP, p. 161), is that use of the DOM marker (Spanish *a*, Hindi *-ko*) decreases, but the use of the homophonous dative marker in ditransitive constructions is maintained. Use of differential subject marking (Hindi ergative *-ne* and dative *-ko*, Spanish dative *a*) also decreases.

The authors discuss several factors that may determine the shape of heritage grammars, including transfer, attrition, incomplete acquisition, and changes in the input (i.e. the language of 1st generation immigrants). Studies on heritage language often don’t seem to provide the data to discern these different factors, the recent paper by Montrul and Sánchez-Walker (2013) being a clear exception, setting a standard for future research. Montrul and Sánchez-Walker report the use of DOM in young and older heritage speakers and native speakers and 1st generation immigrants. Their data shows a clear contrast between the native speakers and the heritage speakers, the 1st generation immigrants sitting in between. If

we take the language of immigrants to resemble the baseline language which served as input to the heritage speaker, which is a very reasonable assumption, these results seem to downplay the effect of attrition (there is no difference between child and adult heritage speakers) and incomplete acquisition (no difference between heritage speakers and the baseline) in this case. This leaves transfer from the dominant language as the main source for the established difference.

Such an effect of transfer can be straightforwardly modeled in López' (2012) analysis of differential object marking in Spanish (with extensions to Hindi, cf. López 2012: Ch. 4). López assumes DOM marking to result from a postsyntactic operation of *Vocabulary Insertion*. Given the right context, a functional syntactic category becomes realized as a differential object marker (*a* in Spanish, *-ko* in Hindi). The same mechanism can presumably be applied to subject case marking. The features that make up the context for the insertion of DOM *a* in Spanish include at least animacy of the object noun, properties of the verb, aspect, and properties of the subject. For Hindi ergative *-ne* these would include at least aspect and features of the verb (transitivity). The realization of these case markers is thus conditioned by a complex of features. It is this complex of features that may be responsible for the difficulty heritage speakers face in the use of differential case marking (cf. Montrul et al. 2012: 168, Montrul and Sánchez-Walker 2013: 112). The transfer effect exists in the fact that the dominant language, English, does not show sensitivity to any of these features in the marking of objects and subjects. English only makes a distinction between indirect object and subject/object by (often) marking the former with a preposition. The English pattern is similar to that found in heritage speakers, who show clear maintenance of indirect object marking only. If we assume transfer from the English pattern to the heritage language, which (gradually) blurs the conditions on the use of the DOM marker, the case marking results fall out naturally without any assumptions about the structural or inherent status of cases.

Thus we see indeed 'a situation where the two factors, dominant-language transfer and incipient change in the input, work together' (BMP, p. 171). This account also ties in neatly with many of the observations provided by BMP. In particular, that morphology is generally the weakest link in heritage grammars (p. 144) and that heritage speakers experience problems with semantics (Section 3.5). Moreover, as López' model locates expression of case marking at the interfaces, it provides additional support for the claim that heritage speakers experience difficulty precisely there (cf. Section 4.2). To advance this kind of analysis, one would have to establish additional facts about the syntactic and semantic behavior of DOM in heritage speakers and of course test speakers with a dominant language different from English.

3 Heritage speakers as single cases

In this final section I want to address a fundamental concern with deriving from the data of heritage speakers generalizations of the kind that structural case is replaced and inherent case is maintained in heritage language. The main problem is that the population of heritage speakers is not exactly a homogenous one and this gets reflected in their linguistic behavior. Consider, as an example, the omission of ergative case by heritage speakers of Hindi reported by Montrul et al. (2012: 159). Heritage speakers leave out the ergative case on average almost 36% of the time, whereas native speakers do so in less than 1%. This looks like a robust difference between the two groups, but individual results sketch a somewhat different picture. Of the 28 heritage speakers tested, 5 never omitted case marking, 6 always omitted it, and the remaining 17 speakers produced 10%–90% omission errors. Likewise, Montrul and Sánchez-Walker (2013: 122) found that adult heritage speakers of Spanish maintained DOM in about 80% of the time on average, but a large group (just over 50%) of these speakers produced DOM 100% of the time (this was 30%–40% with the child heritage speakers). Even though group statistics may come out as significant, I would be hesitant to extract generalizations about a heritage language from these data. Especially, if these generalizations are supposed to have a bearing on theoretical discussions.

Populations used in linguistic research can be placed on a continuum of group homogeneity. On the high end are the native speakers recruited for the average linguistic experiment (the weird people of Henrich et al. 2010) and at the low end we can locate speakers with language-related disorders, like aphasia. Heritage speakers also seem to fall towards the low end of this continuum. They show huge within-group variation not only in language proficiency, but also in age of onset and quantity and quality of input, probably in addition to many other parameters. Each heritage speaker seems to tell his own story. As BMP conjecture, this variation ‘adds to the complexity of heritage language research’ (p. 134). The way forward then is to face this variation head-on. This means that it has to be made central in the design and analysis of experimental studies of heritage language.

The kind of intersubject variability observed in studies with heritage speakers is reminiscent of the variation found in studies with speakers showing a language-related disorder.¹ Participants in the latter kind of studies also each have their own characteristics and background. It has therefore been suggested

¹ I am of course by no means implying that speaking a heritage language amounts to having a language-related disorder!

that it is better not to employ group designs in experimental studies of this population, but to make use of (multiple) single-case studies which allow for a better exploration of intersubject variability (McReynolds and Thompson 1986, see Ruiter et al. 2010 for application of this design in the domain of agrammatism). This bottom-up approach starts at the individual level, establishes a relationship between the individual and the dependent variable of interest (e.g. use of DOM), and then continues to other individuals to see if a generalization can be obtained. This kind of design in combination with methods of analysis (e.g. regression analysis) that allow to simultaneously take into account the influence of several biographic factors that are generally administered in heritage studies, allows for a much more fine-grained analysis of the data than the top-down group model. I would be much more willing to accept a generalization like the structural/inherent case generalization of BMP if it were shown to hold for (a majority of) the individual speakers as well.

Of course, individual characteristics are generally collected in studies with heritage speakers and they figure in the analysis, but often only post-hoc. The main point then is that this biographical information should be used systematically in the analysis of experimental results. Paradoxical as it may sound, putting the individual at center stage will in the end allow for more interesting generalizations that may feed back into linguistic theory.

Examples of this kind can already be glanced from the literature. For instance, Montrul and Sánchez-Walker (2013: 125) show that the omission of DOM marking in Spanish can be related to the exposure to and use of Spanish in daily life. This seems to tie in with the analysis proposed in the previous section. When there is sufficient Spanish input, evidence for the features that provide the context for use of DOM is maintained, giving the English system less space to crawl in. In the domain of DOM, focus on individuals will open up exciting opportunities to study the potential developmental paths of differential case marking systems.

References

- Braine, Martin D. S. 1987. What is learned in acquiring word classes: A step toward an acquisition theory. In B. MacWhinney (ed.), *Mechanisms of language acquisition*, 65–87. Hillsdale, NJ: Erlbaum.
- Corina, D. P., E. K. Gibson, R. Martin, A. Poliakov, J. Brinkley & G. A. Ojemann. 2005. [Dissociation of action and object naming: evidence from cortical stimulation mapping](#). *Human Brain Mapping* 24(1). 1–10.
- Haspelmath, Martin. 2009. Terminology of Case. In A. Malchukov & A. Spencer (eds.), *Oxford Handbook of Case*, 505–517. Oxford: Oxford University Press.

- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663–687.
- Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33(2–3). 61–83.
- Lamers, Monique & Esther Ruigendijk. 2009. Case and Aphasia. In A. Malchukov & A. Spencer (eds.), *Oxford Handbook of Case*, 419–435. Oxford: Oxford University Press.
- van Lier, Eva. 2009. *Parts of speech and dependent clauses: A typological study*. Utrecht: LOT Publications.
- López, Luis. 2012. *Indefinite Objects. Scrambling, Choice Functions, and Differential Marking*. Cambridge MA: The MIT Press.
- McReynolds, Leija V. & Cynthia K. Thompson. 1986. Flexibility of single-subject experimental designs. Part I: Review of the basics of single-subject designs. *Journal of Speech and Hearing Disorders* 51. 194–203.
- Montrul, Silvina, & Melissa Bowles. 2009. Back to basics: Differential object marking under incomplete acquisition in Spanish heritage speakers. *Bilingualism: Language and Cognition* 12(3). 363–383.
- Montrul, Silvina & Noelia Sánchez-Walker. 2013. Differential Object Marking in Child and Adult Spanish Heritage Speakers. *Language Acquisition* 20(2). 109–132.
- Montrul, Silvina, Rakesh Bhatt & Archana Bhatia. 2012. Erosion of case and agreement in Hindi heritage speakers. *Linguistic Approaches to Bilingualism* 2. 141–176.
- Ruiter, Marina B., Herman J. H. Kolk & Toni C. M. Rietveld. 2010. Speaking in ellipses: The effect of a compensatory style of speech on functional communication in chronic agrammatism. *Neuropsychological Rehabilitation: An International Journal* 20(3). 423–458.
- Wasow, Thomas & Jennifer Arnold. 2005. Intuitions in linguistic argumentation. *Lingua* 115. 1481–1496.