

I Congreso Internacional de Investigación y Práctica Profesional en Psicología
XVI Jornadas de Investigación Quinto Encuentro de Investigadores en Psicología
del MERCOSUR. Facultad de Psicología - Universidad de Buenos Aires, Buenos
Aires, 2009.

Efectos del orden gramatical sobre el aprendizaje de un lenguaje artificial.

Tabullo, Angel, Vernis, Sergio, Amil, Ana Belén, Primero, Gerardo, Sánchez, Federico José, Yorio, Alberto, Segura, Enrique y Arismendi, Mariana.

Cita:

Tabullo, Angel, Vernis, Sergio, Amil, Ana Belén, Primero, Gerardo, Sánchez, Federico José, Yorio, Alberto, Segura, Enrique y Arismendi, Mariana (2009). *Efectos del orden gramatical sobre el aprendizaje de un lenguaje artificial. I Congreso Internacional de Investigación y Práctica Profesional en Psicología XVI Jornadas de Investigación Quinto Encuentro de Investigadores en Psicología del MERCOSUR. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires.*

Dirección estable: <https://www.aacademica.org/000-020/141>

ARK: <https://n2t.net/ark:/13683/eYG7/MeW>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

EFFECTOS DEL ORDEN GRAMATICAL SOBRE EL APRENDIZAJE DE UN LENGUAJE ARTIFICIAL

Tabullo, Angel; Vernis, Sergio; Amil, Ana Belén; Primero, Gerardo; Sánchez, Federico José; Yorio, Alberto; Segura, Enrique; Arismendi, Mariana

Facultad de Psicología, Universidad de Buenos Aires

RESUMEN

El primer objetivo de la presente investigación es estudiar los efectos que la frecuencia de uso del orden gramatical produce en el aprendizaje de un lenguaje nuevo. El segundo objetivo es investigar la hipótesis de que los aspectos sintácticos y semánticos del lenguaje son procesados por mecanismos diferentes. Para ello, se diseñaron dos lenguajes artificiales en miniatura. El primero fue denominado gramática frecuente, puesto que presenta un ordenamiento sujeto-objeto-verbo (orden canónico en español). El segundo lenguaje presenta una estructura gramatical OSV, denominada infrecuente. El experimento constó de una fase de entrenamiento, durante la cual el sujeto sólo observaba oraciones correctas, y una de test, en la que debía decidir si las oraciones presentadas contenían violaciones semánticas, sintácticas o ninguna. El grupo GF tuvo un mejor rendimiento en la fase de test, especialmente en la detección de violaciones semánticas. Los resultados apoyan la hipótesis de que un orden gramatical similar al de la primera lengua es más fácil de aprender y sugieren diferencias en el procesamiento de aspectos semánticos y sintácticos durante el aprendizaje de un nuevo lenguaje.

Palabras clave

Lenguaje Gramática Aprendizaje

ABSTRACT

EFFECTS OF WORD ORDER IN ARTIFICIAL LANGUAGE LEARNING

The present experiment has two objectives: to study how word order frequency affects second language learning, and to address the hypothesis that syntactic and semantic aspects of language are processed in different ways. Two miniature artificial languages were designed to that end: a frequent grammar (FG), with a subject-object-verb word order type (canonical order in Spanish), and an infrequent grammar (IG), with an object-subject-verb word order type. The experiment was divided in a training phase, where subjects were exposed to correct sentences only, and a test phase, in which they had to decide if the sentences were correct or incorrect (presenting semantic or syntactic violations). The FG trained group performed better in the test phase, specially in the detection of semantic violations. Our results support the hypothesis that word order types are easier to learn when they are similar to the first language, and suggest differences between semantic and syntactic processing of a new language.

Key words

Language Grammar Learning

INTRODUCCIÓN

Las lenguas suelen disponer, como mínimo, de un orden gramatical básico usado para transmitir, con entonación neutra, una información totalmente nueva. En español, el ordenamiento sujeto-verbo-objeto (SVO) es el de uso más frecuente. Otros órdenes gramaticales, como VSO y VOS, son considerados en nuestra lengua como "órdenes alternativos" (Fernández Soriano, 1993). Existen seis posibles ordenamientos del sujeto, el objeto y el verbo en una oración, y éstos no son igualmente frecuentes en las distintas lenguas alrededor del mundo. Por ejemplo, el ordena-

miento canónico en español (SVO) figura segundo en la lista de frecuencias, presentándose en un 38.7% de los idiomas (Yamamoto, 1999).

Se ha propuesto que la transferencia de los aspectos sintácticos de la lengua materna influye en el aprendizaje de una segunda lengua (Odlin, 1989). El aprendizaje se vería facilitado si ambas lenguas presentan un ordenamiento gramatical similar, y estaría dificultado en el caso contrario (Birdsong & Molis, 2001).

Nuestro primer objetivo consiste en estudiar los efectos que la frecuencia de uso del orden gramatical produce en el aprendizaje de un lenguaje nuevo. Para ello, se diseñaron dos lenguajes artificiales en miniatura. El primero de ellos fue denominado gramática frecuente (GF), puesto que presenta un ordenamiento gramatical SVO (orden canónico en español). El segundo lenguaje presenta una estructura gramatical OSV, gramática denominada infrecuente (GI), no sólo porque rara vez es utilizada en español, sino porque figura última (0.5%) en la lista de frecuencias antes mencionada (op.cit.). Se espera un mejor desempeño de los sujetos hispanoparlantes en el aprendizaje de la gramática frecuente que en el aprendizaje de la gramática infrecuente.

El segundo objetivo es investigar la hipótesis de que los aspectos sintácticos y semánticos del lenguaje son procesados por mecanismos diferentes. Se ha observado que los sujetos encuentran más dificultades en la adquisición de la sintaxis que de la semántica cuando aprenden una segunda lengua (van Hout et al., 2003). Además, mientras que el procesamiento semántico es relativamente similar entre la primera y segunda lengua, el análisis de la información gramatical parece diferir considerablemente entre ambas (Friederici, Steinhauer & Pfeifer, 2002). A nivel neural, de acuerdo con los correlatos electrofisiológicos (Kutas & Van Petten, 1994) y de resonancia magnética funcional (Rüschmeyer et al., 2005), existirían disociaciones en los circuitos cerebrales que se activan en el procesamiento sintáctico y semántico. Esperamos, por lo tanto, que los sujetos entrenados en GF y GI muestren un desempeño desigual en el aprendizaje de los aspectos semánticos y sintácticos del lenguaje. Dado que el grupo entrenado en GF fue expuesto a un orden gramatical similar al de la primera lengua, se espera un mejor desempeño en el aspecto sintáctico, y similar al grupo entrenado en GI en el aspecto semántico.

Para testear ambas hipótesis, se utilizó el procedimiento de gramáticas artificiales, puesto que permite un control experimental de los factores que influyen en la adquisición de una lengua. Además, existe evidencia de que las gramáticas artificiales estarían siendo procesadas como una segunda lengua: estudios realizados con los lenguajes "Brocanto" (Friederici et al., 2002) y "Wernickese" (Newman-Norlund et al., 2006), mostraron que la adquisición de la gramática artificial recluta áreas cerebrales similares a las del procesamiento de la lengua nativa.

MÉTODO

Sujetos: participaron 123 sujetos de ambos sexos, de entre 18 y 48 años de edad, hablantes nativos de español con un nivel de estudios de, al menos, primaria completa. Todos ellos firmaron un consentimiento informado antes de iniciar el estudio.

Procedimiento: los estímulos creados específicamente para esta tarea fueron doce "palabras" bisilábicas, inexistentes en el idioma español. Cinco cumplían función de sustantivos y designaban figuras conocidas de diferentes colores. Otras cinco palabras designaban verbos, que correspondían a diferentes movimientos de las figuras en relación a otra, estática. Los otros dos estímulos eran artículos de una sola sílaba, artículos determinantes para el sujeto y el objeto. Los estímulos eran idénticos en los dos tipos de gramáticas, al igual que su significado.

Las oraciones presentadas tuvieron dos ordenamientos posibles, de acuerdo a la condición experimental asignada. En GF, los estímulos fueron ordenados con la secuencia SVO: Artículo 1 (Fo), Sujeto (cualquiera de los cinco sustantivos), Verbo (cualquiera de los cinco verbos), Artículo 2 (To), Objeto (cualquiera de los cinco sustantivos). La GI presentaba el siguiente orden de estímulos (OSV): Artículo 2, Objeto, Artículo 1, Sujeto, Verbo.

Los sujetos fueron entrenados y testeados individualmente en el aprendizaje de la gramática. Cada sujeto fue asignado aleatoriamente a una de las dos gramáticas artificiales creadas. La tarea estuvo compuesta por dos grandes fases: de entrenamiento y de prueba.

Fase de entrenamiento: se les dio a los sujetos la consigna de observar una serie de escenas y oraciones que las describían, y descubrir la relación entre ellas. Se presentaron 20 escenas de movimientos (**E1**), acompañadas de oraciones en formato audio-visual. Durante esta fase los sujetos podían pasar en cualquier momento a la oración siguiente. En una segunda parte, se intercaló una fase de aprendizaje con feedback (**EF**), en la cual el sujeto tenía que decidir si las oraciones describían la escena en forma correcta o no. Se presentaron 15 ensayos, 5 con frases correctas, 5 con violaciones de tipo sintáctico (inversión de artículos, cambio de verbo por sujeto y de verbo por objeto), y 5 con violación de tipo semántico (verbo incorrecto, sujeto incorrecto, objeto incorrecto), debidamente contrabalanceadas. Se reforzaron las respuestas correctas, y se dio feedback correctivo ante los errores. Luego de esta fase, se presentó otro entrenamiento por exposición (**E2**), con 20 ensayos correctos más.

Fase de prueba: se le dio al sujeto la consigna de clasificar las oraciones según fueran correctas o incorrectas. Se presentaron 45 ensayos con una cantidad equilibrada de frases correctas e incorrectas, y en este último caso, con violaciones tanto de tipo sintáctico como semántico (15 de cada tipo). Se le solicitó que realice la tarea lo más rápidamente posible y tratando de no equivocarse. Durante esta fase, no se suministró feedback.

RESULTADOS

Tiempos de observación: se calculó el promedio del tiempo de observación de las frases durante la fase de exposición en el entrenamiento. El tiempo de observación se comparó con la prueba U-Mann Whitney. Los grupos no se diferenciaron en cuanto al tiempo que invirtieron para pasar de una frase a otra durante E1 ($p=0.371$), pero sí durante E2 ($p=0.015$), siendo los sujetos entrenados en GI los que más tiempo invirtieron en mirar las oraciones.

Porcentaje de aciertos: se utilizó para el análisis un ANOVA de medidas repetidas, con GRUPO como factor intersujetos y TIPO DE ERROR (violaciones semánticas, sintácticas, correctas) como factor intrasujetos. Se halló efecto principal de grupo ($p=0.024$) mostrando el grupo GF un mayor porcentaje de aciertos que GI, mientras que no se encontró efecto de tipo de violación. Sin embargo, hubo efecto de interacción significativo ($p=0.003$) entre ambos factores: GI mostró un menor porcentaje de aciertos en la detección de violaciones semánticas que en las sintácticas ($p=0.001$) y que en las correctas ($p<0.001$), sin diferencia entre estas dos últimas. En cambio, en GF, los sujetos obtuvieron un menor porcentaje de aciertos en la detección de las violaciones sintácticas ($p<0.001$) y semánticas ($p=0.006$) que en las correctas. El grupo GF tuvo un mejor rendimiento que GI en la detección de oraciones correctas ($p=0.045$) y en la de violaciones semánticas ($p=0.001$). Ambos grupos no se diferenciaron significativamente en la detección de errores sintácticos.

Tiempos de respuesta: se realizó un ANOVA grupo X tipo de error con los datos normalizados; se encontró un efecto principal de tipo de error ($p<0.001$): los sujetos respondieron más rápidamente en las oraciones correctas que en las violaciones ($p<0.001$). Los tiempos de respuesta en las violaciones semánticas fueron menores que en las sintácticas ($p<0.001$). Por su parte, también se encontró un efecto principal de grupo ($p=0.016$): las respuestas de los sujetos del grupo GF fueron más rápidas que las del grupo GI. No se encontró interacción.

DISCUSIÓN

Los resultados obtenidos apoyan nuestra primera predicción. Los sujetos entrenados en GI no sólo necesitaron invertir más tiempo observando las oraciones, sino que también cometieron más errores y respondieron más lentamente en la fase de prueba. Esto apoya la hipótesis de que el aprendizaje de una segunda lengua se facilita cuando su ordenamiento gramatical coincide con el de la primera lengua, y se dificulta en el caso contrario.

Nuestra segunda predicción respecto de la segregación en el procesamiento de los aspectos semánticos y sintácticos del lenguaje fue parcialmente confirmada. Observamos desigualdades en la detección de errores semánticos y sintácticos, aunque no en el sentido esperado. Los sujetos entrenados en GF tuvieron mejor desempeño en los aspectos semánticos, pero no se encontraron diferencias en los aspectos sintácticos de la gramática.

¿El rendimiento inferior de los sujetos entrenados en GI es congruente con estudios previos. Simulaciones computacionales realizadas con redes neurales (Lupyan & Christiansen, 2002) mostraron que el desempeño en el aprendizaje de roles gramaticales varía según el orden gramatical presentado. Las redes fueron más eficientes en los ordenamientos gramaticales con mayor frecuencia de uso alrededor del mundo. En otro estudio realizado con niños (Johnston, Blatchley, Olness, 1990) se observó que un lenguaje artificial SOV resultó más fácil de aprender que uno VSO. ¿Los autores interpretan estos efectos del orden gramatical como el reflejo de limitaciones cognitivas para el aprendizaje de estructuras secuenciales, las cuales conducirían al sesgo en la distribución de los órdenes gramaticales. Esta hipótesis puede ser complementaria o alternativa a la de la transferencia sintáctica, que también explicaría nuestros resultados.

El hecho de que los grupos en nuestra investigación mostraran diferencias de desempeño en los aspectos semánticos pero no en los sintácticos puede explicarse apelando a una hipótesis alternativa. Para detectar las violaciones semánticas, los sujetos debían comprender los roles gramaticales dentro de las oraciones (S,O,V) y compararlos con lo que mostraban las escenas. Estos roles serían más difíciles de identificar en el caso de la gramática infrecuente, afectando el desempeño. Para detectar las violaciones sintácticas, en cambio, los sujetos podrían haber recurrido a la probabilidad de transición entre palabras. Cuando se presenta una anomalía sintáctica, los sujetos observan aparecer en forma consecutiva palabras que nunca estuvieron juntas durante el entrenamiento, es decir, cuya probabilidad transicional es nula. Dado que esta información es independiente de los roles gramaticales, está igualmente disponible en ambos lenguajes, y es por eso que los sujetos no se diferenciarían significativamente en la detección de estos errores. Investigaciones recientes apoyan la participación de mecanismos estadísticos en la adquisición del lenguaje, en particular, de la sintaxis (Saffran, 2002).

En síntesis, nuestro experimento mostró que el orden gramatical tiene efectos en el aprendizaje de un lenguaje nuevo, requiriéndose experimentos adicionales para determinar si dichos efectos se deben a la transferencia generada por la lengua materna o dependen del orden gramatical en sí mismo. Se observaron, además, disociaciones en el desempeño en aspectos semánticos y sintácticos de los lenguajes, atribuibles a fenómenos de tipo estadístico.

BIBLIOGRAFÍA

- BIRDSOING, D. & MOLIS, M. (2001). On the evidence for maturational effects in second language acquisition. *Journal of Memory and Language*, 44, 235-249.
- CHRISTIANSEN, M.H. (1994). Infinite languages, finite minds: Connectionism, learning and linguistic structure. Manuscrito no publicado.
- FERNANDEZ SORIANO, O. (1993). Sobre el orden de palabras en español. *Dicenda. Cuadernos de Filología Hispánica*, 11, 113-152.
- FRIEDERICI, A. et al. (2002). Brain signatures of artificial language processing: evidence challenging the "critical period" hypothesis. *Proceedings of the National Academy of Science USA*, 99, 529-534.
- KUTAS, M.; VAN PETTEN, C. (1994). Psycholinguistics electrified. Event related brain potential investigations. En M. A. Gernsbacher, (ed.), *Handbook of psycholinguistics* (pp.83-143). San Diego: Academic Press.
- LUPYAN, G. & CHRISTIANSEN, M.H. (2002). Case, Word Order and Language Learnability: Insights from Connectionist Modeling. *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, 596-601.
- NEWMAN-NORLUND, R.D.; FREY, S.H.; PETITTO, L.A. & GRAFTON, S.T. (2006). Anatomical Substrates of Visual and Auditory Miniature Second-language Learning. *Journal of Cognitive Neuroscience*, 18, 1984-1997.
- ODLIN, T. (1989). *Language Transfer: Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- RÜSCHEMEYER, S.; FIEBACH, C.; KEMPE, V.; FRIEDERICI, A. (2005). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Human brain mapping*, 25, 266-86.
- SAFFRAN, J.R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, 47, 172-196.
- VAN HOUT, R., et al. (eds.) (2003) *The Lexicon-Syntax Interface in Second Language Acquisition*, John Benjamins Publishing Company.
- YAMAMOTO, H. (1999). A survey of areal distribution of word order around the world. *Sprachtypologie und Universalienforschung*, 52, 64-77.